

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2002, 19:44:02 ; Search time 1535.11 Seconds

(Without alignments)
236.425 Million cell updates/sec

Title: US-09-802-518-1

Perfect score: 1 tgactgtgacgttcgagatga 22

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*
1: gb_da.*
2: gb_hhg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*
15: gb_da.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_om.*
20: em_or.*
21: em_ov.*
22: em_pat.*
23: em_ph.*
24: em_pl.*
25: em_ro.*
26: em_sts.*
27: em_sy.*
28: em_un.*
29: em_vl.*
30: em_hhgo_hum.*
31: em_hhgo_inv.*
32: em_hhgo_rod.*
33: em_hhg_hum.*
34: em_hhg_inv.*
35: em_hhg_rod.*
36: em_hhg_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	22	100.0	22	6	AX036945	AX036945 Sequence
2	22	100.0	22	6	AX046593	AX046593 Sequence
3	22	100.0	22	6	AX083675	AX083675 Sequence
4	22	100.0	22	6	AX135650	AX135650 Sequence
5	22	100.0	22	6	AX148636	AX148636 Sequence
6	21	95.5	22	6	AX083681	AX083681 Sequence
7	21	95.5	22	6	AX148642	AX148642 Sequence
8	20.4	92.7	22	6	AX148608	AX148608 Sequence
9	20.4	92.7	22	6	AX036946	AX036946 Sequence
10	20.4	92.7	22	6	AX083676	AX083676 Sequence
11	20.4	92.7	22	6	AX083678	AX083678 Sequence
12	20.4	92.7	22	6	AX148637	AX148637 Sequence
13	20.4	92.7	22	6	AX148639	AX148639 Sequence
14	20.2	91.8	22	6	AX148643	AX148643 Sequence
15	20	90.9	22	6	AX083682	AX083682 Sequence
16	20	90.9	22	6	AX174913	AX174913 Sequence
17	19.4	88.2	22	6	AX083680	AX083680 Sequence
18	19.4	88.2	22	6	AX148641	AX148641 Sequence
19	18.8	85.5	22	6	AX148607	AX148607 Sequence
20	18.8	85.5	22	6	AX148609	AX148609 Sequence
21	18.8	85.5	22	6	AX148616	AX148616 Sequence
22	18.8	85.5	22	6	AX036944	AX036944 Sequence
23	18.8	85.5	22	6	AX036952	AX036952 Sequence
24	18.8	85.5	22	6	AX135651	AX135651 Sequence
25	18.8	85.5	22	6	AX148644	AX148644 Sequence
26	18.8	85.5	22	6	AX148645	AX148645 Sequence
27	17.8	80.9	165337	2	AC027442	AC027442 Homo sapi
28	17.8	80.9	167237	9	AC009738	AC009738 Homo sapi
29	17.8	80.9	177371	2	AC092138	AC092138 Homo sapi
30	17.8	80.9	201214	2	AC074012	AC074012 Homo sapi
31	17.8	80.9	204992	2	AC024934	AC024934 Homo sapi
32	17.2	78.2	22	6	AR148610	AR148610 Sequence
33	17.2	78.2	22	6	AX135652	AX135652 Sequence
34	17.2	78.2	127	6	AC088	272011 A.gambiae s
35	17.2	78.2	552	8	RICSHNE04	D85048 Rice retro
36	17.2	78.2	12945	3	CEP42610	Z48330 Caenorhabdi
37	17.2	78.2	85219	9	AC073520	AC073520 Homo sapi
38	17.2	78.2	110000	2	AC091528	Continuation (3 of
39	17.2	78.2	128216	9	HS01319M	AL079341 Human DNA
40	17.2	78.2	169271	2	AC079336	AC079336 Homo sapi
41	17.2	78.2	169277	2	AC078840	AC078840 Homo sapi
42	17.2	78.2	183827	9	AC005899	AC005899 Homo sapi
43	17.2	78.2	185784	2	AC009719	AC009719 Homo sapi
44	17.2	78.2	265040	2	AC091363	AC091363 Rattus no
45	16.8	76.4	2040	3	AF115507	AF115507 Leishmani

ALIGNMENTS

RESULT 1
LOCUS AX036945 22 bp DNA PAT 16-NOV-2000
DEFINITION Sequence 2 from Patent FR2790955.
ACCESSION AX036945
VERSION AX036945.1 GI:11226373

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
JOURNAL
ASSIST PUBL HOPITAUX DE PARIS (FR)
Location/Qualifiers
1..22
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Oligodesoxynucleotide"

BASE COUNT 6 a 3 c 7 g 6 t

ORIGIN

Query Match 100.0%; Score 22; DB 6; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgactgtgaacgttcgagatga 22
|||||
Db 1 TGACTGTGAACGTCGAGATGA 22

RESULT 2

AX046993

LOCUS AX046993 22 bp DNA

DEFINITION Sequence 2 from Patent WO0067787.

ACCESSION AX046993

VERSION AX046993.1 GI:11876420

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

SOURCE

1.22

/organism="synthetic construct"

/db_xref="taxon:32630"

/note="phosphorothioate-modified synthetic

oligodeoxynucleotide"

BASE COUNT

ORIGIN

6 a 3 c 7 g 6 t

Query Match

Best Local Similarity

Matches

22; Conservative

0; Mismatches

0; Indels

0; Gaps

0;

OY 1 tgactgtgaacgttcgagatga 22

|||||

Db 1 TGACTGTGAACGTCGAGATGA 22

RESULT 3

AX083675

LOCUS AX083675 22 bp DNA

DEFINITION Sequence 1 from Patent WO0112223.

ACCESSION AX083675

VERSION AX083675.1 GI:13185407

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

SOURCE

1.22

/organism="synthetic construct"

/db_xref="taxon:32630"

/note="Synthetic construct"

BASE COUNT

ORIGIN

6 a 3 c 7 g 6 t

Query Match

Best Local Similarity

100.0%; Score 22; DB 6; Length 22;

100.0%; Pred. No. 0.13;

Matches 22; Conservative

0; Mismatches

0; Indels

0; Gaps

0;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgactgtgaacgttcgagatga 22
|||||
Db 1 TGACTGTGAACGTCGAGATGA 22

RESULT 4

AX135650

LOCUS AX135650 22 bp DNA

DEFINITION Sequence 21 from Patent WO0132877.

ACCESSION AX135650

VERSION AX135650.1 GI:14271920

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

SOURCE

1.22

/organism="synthetic construct"

/db_xref="taxon:32630"

/note="Cpg oligonucleotide"

BASE COUNT

ORIGIN

6 a 3 c 7 g 6 t

Query Match

Best Local Similarity

Matches

22; Conservative

0; Mismatches

0; Indels

0; Gaps

0;

OY 1 tgactgtgaacgttcgagatga 22

|||||

Db 1 TGACTGTGAACGTCGAGATGA 22

RESULT 5

AX148636

LOCUS AX148636 22 bp DNA

DEFINITION Sequence 1 from Patent WO0135991.

ACCESSION AX148636

VERSION AX148636.1 GI:14347254

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

SOURCE

1.22

/organism="synthetic construct"

/db_xref="taxon:32630"

/note="Synthetic construct"

BASE COUNT

ORIGIN

6 a 3 c 7 g 6 t

Query Match

Best Local Similarity

100.0%; Score 22; DB 6; Length 22;

100.0%; Pred. No. 0.13;

Matches 22; Conservative

0; Mismatches

0; Indels

0; Gaps

0;

RESULT 6
AX083681 22 bp DNA PAT 28-FEB-2001
LOCUS AX083681
DEFINITION Sequence 7 from Patent WO0112223.
ACCESSION AX083681
VERSION AX083681.1 GI:13185413
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (bases 1 to 22)
AUTHORS van Nest,G.
TITLE Methods of modulating an immune response using immunostimulatory s
JOURNAL Sequences and compositions for use therein
Patent: WO 0112223-A 7 22-FEB-2001;
FEATURES Location/Qualifiers
source 1..22
/organism="synthetic construct"
/db_xref="taxon:32630"
modified_base 11
/note="5-bromocytosine"
BASE COUNT 6 a 2 c 7 g 6 t 1 others
ORIGIN

Query Match 95.5%; Score 21; DB 6; Length 22;
Best Local Similarity 95.5%; Pred. No. 0.49;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 tgactgtgaacgttcgagatga 22
|||||
Db 1 TGA CTGTGA NGTTCGAGATGA 22

RESULT 7
AX148642 22 bp DNA PAT 08-JUN-2001
LOCUS AX148642
DEFINITION Sequence 7 from Patent WO0135991.
ACCESSION AX148642
VERSION AX148642.1 GI:14347260
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (bases 1 to 22)
AUTHORS Tuck,S. and van Nest,G.
TITLE Immunomodulatory compositions containing an immunostimulatory
JOURNAL Sequence linked to antigen and methods of use thereof
Patent: WO 0135991-A 7 25-MAY-2001;
FEATURES Location/Qualifiers
source 1..22
/organism="synthetic construct"
/db_xref="taxon:32630"
modified_base 11
/note="synthetic construct"
/note="5-bromocytosine"
BASE COUNT 6 a 2 c 7 g 6 t 1 others
ORIGIN

Query Match 95.5%; Score 21; DB 6; Length 22;
Best Local Similarity 95.5%; Pred. No. 0.49;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 tgactgtgaacgttcgagatga 22
|||||
Db 1 TGA CTGTGA NGTTCGAGATGA 22

RESULT 8
ARI48608 22 bp DNA PAT 08-AUG-2001
LOCUS ARI48608
DEFINITION Sequence 2 from patent US 6225292.
ACCESSION ARI48608
VERSION ARI48608.1 GI:15112698
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 22)
AUTHORS Raz,E. and Roman,M.
TITLE Inhibitors of DNA Immunostimulatory sequence activity
JOURNAL Patent: US 6225292-A 2 01-MAY-2001;
FEATURES Location/Qualifiers
source 1..22
/organism="unknown"
BASE COUNT 7 a 2 c 7 g 6 t
ORIGIN

Query Match 92.7%; Score 20.4; DB 6; Length 22;
Best Local Similarity 95.5%; Pred. No. 1.1;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 tgactgtgaacgttcgagatga 22
|||||
Db 1 TGA CTGTGA NGTTCGAGATGA 22

RESULT 9
AX036946 22 bp DNA PAT 16-NOV-2000
LOCUS AX036946
DEFINITION Sequence 3 from Patent FR2790955.
ACCESSION AX036946
VERSION AX036946.1 GI:11226374
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (bases 1 to 22)
AUTHORS Carpentier,A.
JOURNAL Patent: FR 2790955-A 3 22-SEP-2000;
FEATURES Location/Qualifiers
source 1..22
/organism="synthetic construct"
/db_xref="taxon:32630"
modified_base 11
/note="oligodeoxynucleotide"
BASE COUNT 6 a 4 c 6 g 6 t
ORIGIN

Query Match 92.7%; Score 20.4; DB 6; Length 22;
Best Local Similarity 95.5%; Pred. No. 1.1;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 tgactgtgaacgttcgagatga 22
|||||
Db 1 TGA CTGTGA NGTTCGAGATGA 22

RESULT 10
AX083676 22 bp DNA PAT 28-FEB-2001
LOCUS AX083676
DEFINITION Sequence 2 from Patent WO0112223.
ACCESSION AX083676
VERSION AX083676.1 GI:13185408
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.

REFERENCE 1 (bases 1 to 22)
AUTHORS Van Nest,G.
TITLE Methods of modulating an immune response using immunostimulatory s
JOURNML Patent: WO 0112223-A 2 22-FEB-2001;
FEATURES Dynavax Technologies Corporation (US)
SOURCE Location/Qualifiers
1. .22
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Synthetic construct"
BASE COUNT 6 a 4 c 7 g 5 t
ORIGIN

Query Match 92.7%; Score 20.4; DB 6; Length 22;
Best Local Similarity 95.5%; Pred. No. 1.1;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 tgactgtgaacgttcgagatga 22
|||||
Db 1 TGACCGTGAACGCTTCGAGATGA 22

RESULT 11
AX083678 22 bp DNA PAT 28-FEB-2001
LOCUS AX083678
DEFINITION Sequence 4 from Patent WO0112223.
ACCESSION AX083678
VERSION AX083678.1 GI:13185410
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 22)
AUTHORS van Nest,G.
TITLE Methods of modulating an immune response using immunostimulatory s
JOURNML Patent: WO 0112223-A 4 22-FEB-2001;
FEATURES Dynavax Technologies Corporation (US)
SOURCE Location/Qualifiers
1. .22
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Synthetic construct"
BASE COUNT 6 a 4 c 6 g 6 t
ORIGIN

Query Match 92.7%; Score 20.4; DB 6; Length 22;
Best Local Similarity 95.5%; Pred. No. 1.1;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 tgactgtgaacgttcgagatga 22
|||||
Db 1 TGACTGTGAACGCTTCACATGA 22

RESULT 12
AX148637 22 bp DNA PAT 08-JUN-2001
LOCUS AX148637
DEFINITION Sequence 2 from Patent WO0135991.
ACCESSION AX148637
VERSION AX148637.1 GI:14347255
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 22)
AUTHORS Tuck,S. and Van Nest,G.
TITLE Immunomodulatory compositions containing an immunostimulatory
JOURNML Patent: WO 0135991-A 2 25-MAY-2001;

Dynavax Technologies Corporation (US)
FEATURES Location/Qualifiers
1. .22
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Synthetic construct"
BASE COUNT 6 a 4 c 7 g 5 t
ORIGIN

Query Match 92.7%; Score 20.4; DB 6; Length 22;
Best Local Similarity 95.5%; Pred. No. 1.1;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 tgactgtgaacgttcgagatga 22
|||||
Db 1 TGACCGTGAACGCTTCGAGATGA 22

RESULT 13
AX148639 22 bp DNA PAT 08-JUN-2001
LOCUS AX148639
DEFINITION Sequence 4 from Patent WO0135991.
ACCESSION AX148639
VERSION AX148639.1 GI:14347257
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 22)
AUTHORS Tuck,S. and Van Nest,G.
TITLE Immunomodulatory compositions containing an immunostimulatory
JOURNML Patent: WO 0135991-A 4 25-MAY-2001;
FEATURES Dynavax Technologies Corporation (US)
SOURCE Location/Qualifiers
1. .22
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Synthetic construct"
BASE COUNT 6 a 4 c 6 g 6 t
ORIGIN

Query Match 92.7%; Score 20.4; DB 6; Length 22;
Best Local Similarity 95.5%; Pred. No. 1.1;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 tgactgtgaacgttcgagatga 22
|||||
Db 1 TGACTGTGAACGCTTCACATGA 22

RESULT 14
AX148643 22 bp DNA PAT 08-JUN-2001
LOCUS AX148643
DEFINITION Sequence 8 from Patent WO0135991.
ACCESSION AX148643
VERSION AX148643.1 GI:14347261
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 22)
AUTHORS Tuck,S. and Van Nest,G.
TITLE Immunomodulatory compositions containing an immunostimulatory
JOURNML Patent: WO 0135991-A 8 25-MAY-2001;
FEATURES Dynavax Technologies Corporation (US)
SOURCE Location/Qualifiers
1. .22
/organism="synthetic construct"
/db_xref="taxon:32630"

modified_base 11 /note="synthetic construct"

/note="5-bromocytosine"

BASE COUNT 6 a 1 c 7 g 6 t 2 others

ORIGIN

Query Match 91.8%; Score 20.2; DB 6; Length 22;

Best Local Similarity 90.9%; Pred. No. 1.4;

Matches 20; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 tgactgtgaacgttcgagatga 22

Db 1 TGA CTGTGA NGTTBGAGATGA 22

RESULT 15

AX083682

LOCUS

AX083682

DEFINITION

AX083682

ACCESSION

AX083682.1

VERSION

AX083682.1

KEYWORDS

SYNTHETIC CONSTRUCT

ORGANISM

synthetic construct

REFERENCE

1 (bases 1 to 22)

AUTHORS

van Nest, G.

TITLE

Methods of modulating an immune response using immunostimulatory s

JOURNAL

Patent: WO 012223-A 8 22-FEB-2001;

FEATURES

Location/Qualifiers

1..22

source

modified_base

11

/db_xref="taxon:32630"

modified_base

15

/note="5-bromocytosine"

modified_base

15

/note="5-bromocytosine"

BASE COUNT

6 a 1 c 7 g 6 t 2 others

ORIGIN

Query Match 90.9%; Score 20; DB 6; Length 22;

Best Local Similarity 90.9%; Pred. No. 1.8;

Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 tgactgtgaacgttcgagatga 22

Db 1 TGA CTGTGA NGTTNGAGATGA 22

Search completed: March 11, 2002, 20:55:45
Job time: 4303 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: March 11, 2002, 19:50:17 ; Search time 227.21 Seconds
(without alignments)
83.012 Million cell updates/sec

Title: US-09-802-518-1

Sequence: 1 tgactgtgacgttcgagatga 22

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 930621 segs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_1101.*
1: /SIDSL/gcgdata/geneseq/geneseqn/NA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseqn/NA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseqn/NA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseqn/NA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseqn/NA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseqn/NA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseqn/NA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseqn/NA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseqn/NA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseqn/NA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseqn/NA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseqn/NA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseqn/NA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseqn/NA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseqn/NA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseqn/NA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseqn/NA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseqn/NA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseqn/NA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseqn/NA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseqn/NA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseqn/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	100.0	22	19	AAV32079
2	22	100.0	22	20	AAV36624
3	22	100.0	22	20	AAV80097
4	22	100.0	22	20	AAV80102
5	22	100.0	22	20	AAV80103
6	22	100.0	22	21	AAV64051
7	22	100.0	22	21	AAV6253
8	22	100.0	22	21	AAV90458
9	22	100.0	22	21	AAV14467
10	22	100.0	22	21	AAV38065
11	22	100.0	22	21	AAV38071

12	22	100.0	22	21	AAV38072	Immunostimulatory
13	22	100.0	22	21	AAV38076	Immunomodulatory o
14	22	100.0	22	22	AAH42533	Phosphorothioate b
15	22	100.0	22	22	AAH73439	Immunomodulatory n
16	22	100.0	22	22	AAH44109	5' terminal NH2 gr
17	22	100.0	22	22	AAH41573	Immunostimulatory
18	22	100.0	22	22	AAH20403	Cpg motif cental
19	22	100.0	22	22	AAH77040	Immunomodulatory D
20	22	100.0	22	22	AAH29600	Cholera toxin immu
21	22	100.0	22	22	AAH82107	CG motif and CFA c
22	22	100.0	22	22	AAH92377	Immunostimulatory o
23	22	100.0	22	22	AAH77046	Immunomodulatory o
24	21	95.5	22	21	AAH55880	Immunomodulatory o
25	21	95.5	22	22	AAH41579	Immunostimulatory
26	20.4	92.7	22	20	AAV80095	Oligo used in expe
27	20.4	92.7	22	20	AAV80096	Immunomodulatory o
28	20.4	92.7	22	20	AAV80099	Immunomodulatory o
29	20.4	92.7	22	20	AAV80101	Immunomodulatory o
30	20.4	92.7	22	21	AAV6254	Sequence of a stab
31	20.4	92.7	22	21	AAV38066	Immunostimulatory
32	20.4	92.7	22	21	AAV38068	Immunostimulatory
33	20.4	92.7	22	21	AAV38070	Immunostimulatory
34	20.4	92.7	22	22	AAH42534	Phosphorothioate b
35	20.4	92.7	22	22	AAH73440	Immunomodulatory n
36	20.4	92.7	22	22	AAH41574	Immunostimulatory
37	20.4	92.7	22	22	AAH41576	Immunostimulatory
38	20.4	92.7	22	22	AAH77041	Immunostimulatory
39	20.4	92.7	22	22	AAH77043	Immunostimulatory
40	20.4	92.7	22	22	AAH77047	Immunostimulatory
41	20.2	91.8	22	22	AAH41580	Immunostimulatory
42	20	90.9	22	21	AAH55881	Immunostimulatory o
43	19.6	89.1	22	22	AAH77045	Immunostimulatory
44	19.4	88.2	22	21	AAH55877	Immunostimulatory o
45	19.4	88.2	22	22	AAH41578	Immunostimulatory

ALIGNMENTS

RESULT 1	AAV32079	standard; DNA; 22 BP.
ID	AAV32079	
XX	AAV32079;	
AC		
XX		
DT	09-SEP-1998	(first entry)
XX		
DE		
XX		
KW	DY1018; beta-gal; ISS-PN/IMM; antigen; immune response; antibody;	
KW	Immunisation; anapylaxis; Ige; retinopathies; ss.	
XX		
OS	synthetic.	
XX		
FH	Key	Location/Qualifiers
FT	modified_base	1..22
FT		/*tag= a
FT		/note="phosphothioate backbone"
XX		
PN	WO9816247-A1.	
XX		
PD	23-APR-1998.	
XX		
PF	09-OCT-1997;	97MO-US19004.
XX		
PR	11-OCT-1996;	96US-0028118.
XX		
PA	(REGC) UNIV CALIFORNIA.	
XX		
PI	Carson DA, Raz E, Roman M;	
XX		
DR	WPI, 1998-261028/23.	
XX		

Tue Mar 12 10:44:32 2002

PT New immunomodulatory compositions - comprising an antigen conjugated to a polynucleotide that contains an immunostimulatory sequence

Example 1: Page 36; 69pp; English.

XX This is the nucleotide sequence of DY1018, which is conjugated to
CC beta-gal to form ISS-PN/IMM, comprising an immunomodulatory molecule
CC (IMM), which comprises an antigen conjugated to a polynucleotide
CC (PN) that contains at least one immunostimulatory nucleotide sequence
CC (ISS). The conjugate synergistically boost the magnitude of the host
CC immune response against an antigen to a level greater than the host
CC immune response to either the IMM, antigen or ISS-PN alone. These
CC responses to ISS-PN/IMM conjugates boost both humoral (antibody) and cellular
CC the important early phase of the host immune response to an antigen.
CC The ISS-PN/IMM conjugates boost the host. Thus, use of the method to
CC (Th1 type) immune responsiveness of a host to subsequent challenge by a
CC boost the immune responsiveness of a host to subsequent challenge by a
CC sensitizing antigen without immunisation avoids the risk of
CC Th2-mediated, immunisation-induced anaphylaxis by suppressing IgE
CC production in response to the antigen challenge. The conjugates can
CC also be used to combat pathogenic infection and to stimulate
CC therapeutic angiogenesis to treat conditions in which localised blood
CC flow plays a significant etiological role, e.g. retinopathies.

Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;

Query Match 100.0%; Score 22; DB 19; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.034; Mismatches 0; Indels 0; Gaps 0;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 tgactgtgaacgttcgagatga 22
1 tgactgtgaacgttcgagatga 22

DB

RESULT 2

AA36624 standard; DNA: 22 BP.

XX AAX36624:

XX 09-JUL-1999 (first entry)

XX ISS-ODN DY1018 nucleotide sequence.

XX Antigen-stimulated inflammation; immunostimulatory oligonucleotide;

XX granulocyte-mediated tissue inflammation; Th2 type immune response;

XX immune responsiveness modulation; idiopathic hypersensitivity; ss;

XX cutaneous basophil hypersensitivity; ISS-ODN; asthma; nasal polyposis;

XX allergic rhinitis; atopic dermatitis; allergic conjunctivitis;

XX eosinophilic fasciitis; therapy; ss.

XX Synthetic.

XX WO9911275-A2.

XX 11-MAR-1999.

XX 04-SEP-1998; 98WO-US18382.

XX 05-SEP-1997; 97US-0927120.

XX (REGC) UNTV CALIFORNIA.

XX Ray E;

XX WPI; 1999-312404/26.

XX Reducing antigen-stimulated granulocyte-mediated inflammation

XX Example 2; Page 30; 69pp; English.

XX

XX

PS

CC This is the ISS-ODN DY1018 nucleotide sequence.
CC The invention relates to a method for preventing or reducing in a
CC antigen-stimulated, granulocyte-mediated tissue inflammation in a
CC by administering an immunostimulatory oligonucleotide (ISS-ODN), where
CC (a) reduction in, or the absence of, a Th2 type immune response is
CC measured; or (b) there is a reduction or absence of other clinical signs
CC of inflammation in the host after antigen challenge. The method is used
CC to reduce or suppress granulocyte-mediated inflammation in a host tissue,
CC and to modulate the host's immune responsiveness to an antigen,
CC particularly where the subject suffers from asthma, nasal polyposis,
CC allergic rhinitis, atopic dermatitis, allergic conjunctivitis, or
CC eosinophilic fasciitis, idiopathic hypersensitivity, or treatment by
CC cutaneous basophil hypersensitivity. Unlike prior art treatment method,
CC antigen immunisation, the method is an antigen-independent method,
CC and avoids host production of both interleukin-4 (IL-4), which carries
CC risk of anaphylaxis, and IL-5 which actually encourages granulocyte
CC adhesion to endothelia.

Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;

Query Match 100.0%; Score 22; DB 20; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.034; Mismatches 0; Indels 0; Gaps 0;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 tgactgtgaacgttcgagatga 22
1 tgactgtgaacgttcgagatga 22

DB

RESULT 3

AA36624 standard; DNA: 22 BP.

XX AAX36624:

XX 12-MAR-1999 (first entry)

XX Immunomodulatory oligo comprising an ISS sequence.

XX Immunomodulatory oligo comprising an ISS sequence; immune regulation;

XX immunomodulatory; immunostimulatory; octanucleotide; papillomavirus;

XX ISS; cancer; allergy; asthma; hepatitis B infection; papillomavirus; ss;

XX human immunodeficiency virus; influenza; herpes; M. tuberculosis; Schistosoma.

XX B. pertussis; malaria; plasmodia; leishmania; Trypanosoma;

XX Synthetic.

XX WO9855495-A2.

XX 10-DEC-1998.

XX 05-JUN-1998; 98WO-US11578.

XX 06-JUN-1997; 97US-0048793.

XX (DYNA-) DYNAXX TECHNOLOGIES CORP.

XX Dina D, Roman M, Schwartz D;

XX WPI; 1999-059898/05.

XX Immunostimulatory oligonucleotides regulate the immune system - and

XX contain an immune-stimulating octanucleotide sequence; for treating

XX cancer, allergic and infectious diseases

XX Claim 5; Page 29; 63pp; English.

XX The invention relates to immunomodulatory oligonucleotides that comprise

XX at least 1 immunostimulatory octanucleotide sequence (ISS) where the IS;

XX sequences are selected from the group consisting of AACGTCG, AACGTCG,

XX GACGTCG, and GACGTCG. The immunomodulatory sequences are used to treat

XX patients needing immune regulation, such as those suffering from cancer,

XX an allergic disease and asthma. They are also used to prevent infection

CC

CC diseases such as influenza, herpes, hepatitis B, human immunodeficiency
 CC and papillomavirus, Hemophilus influenza, Mycobacterium tuberculosis and
 CC Bordetella pertussis, malarial plasmodia, Leishmania, Trypanosoma and
 CC Schistosoma. The immunomodulatory sequences are used to screen for human
 CC immunostimulatory activity by incubating macrophage cells and the
 CC oligonucleotide; and determining the relative amount of Th1-biased
 CC cytokines in the supernatant. Sequences AAV80096 to AAV80103 represent
 CC specific claimed examples of such immunomodulatory oligonucleotides.
 XX
 SO Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;

Query Match 100.0%; Score 22; DB 20; Length 22;
 Best Local Similarity 100.0%; Pred. No. 0.034;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgactgtgaacgttcgagatga 22
 |||||
 Db 1 tgactgtgaacgttcgagatga 22

RESULT 4

AAV80102 standard; DNA; 22 BP.

AC AAV80102;

DT 12-MAR-1999 (first entry)

DE Immunomodulatory oligo comprising an ISS sequence.

XX Immunomodulatory; immunostimulatory; octanucleotide; immune regulation;
 KW ISS; cancer; allergy; asthma; hepatitis B infection; papillomavirus;
 KW human immunodeficiency virus; influenza; herpes; M. tuberculosis; ss;
 KW B. pertussis; malaria; plasmodia; Leishmania; Trypanosoma; Schistosoma.
 XX
 OS Synthetic.

XX Key modified_base 11

FT Location/Qualifiers

FT /*tag= a

FT /note= "5-bromocytosine"

XX WO9855495-A2.

PD 10-DEC-1998.

XX 05-JUN-1998; 98WO-US11578.

XX 06-JUN-1997; 97US-0048793.

XX (DYNA-) DYNAVAX TECHNOLOGIES CORP.

XX Dina D, Roman M, Schwartz D;

XX WPI: 1999-059898/05.

PT Immunostimulatory oligonucleotides regulate the immune system - and
 PT contain an immune-stimulating octanucleotide sequence; for treating
 PT cancer, allergic and infectious diseases
 XX
 PS Claim 23; Page 30; 63pp; English.

XX The invention relates to immunomodulatory oligonucleotides that comprise
 CC at least 1 immunostimulatory octanucleotide sequence (ISS) where the ISS
 CC sequences are selected from the group consisting of AACGTTCC, AACGTTCC,
 CC GACGTTCC, and GACGTTCC. The immunomodulatory sequences are used to treat
 CC patients needing immune regulation, such as those suffering from cancer,
 CC an allergic disease and asthma. They are also used to prevent infectious
 CC diseases such as influenza, herpes, hepatitis B, human immunodeficiency
 CC and papillomavirus, Hemophilus influenza, Mycobacterium tuberculosis and
 CC Bordetella pertussis, malarial plasmodia, Leishmania, Trypanosoma and
 CC Schistosoma. The immunomodulatory sequences are used to screen for human

CC immunostimulatory activity by incubating macrophage cells and the
 CC oligonucleotide; and determining the relative amount of Th1-biased
 CC cytokines in the supernatant. Sequences AAV80096 to AAV80103 represent
 CC specific claimed examples of such immunomodulatory oligonucleotides.
 XX
 SO Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;

Query Match 100.0%; Score 22; DB 20; Length 22;
 Best Local Similarity 100.0%; Pred. No. 0.034;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgactgtgaacgttcgagatga 22
 |||||
 Db 1 tgactgtgaacgttcgagatga 22

RESULT 5

AAV80103 standard; DNA; 22 BP.

AC AAV80103;

DT 12-MAR-1999 (first entry)

DE Immunomodulatory oligo comprising an ISS sequence.

XX Immunomodulatory; immunostimulatory; octanucleotide; immune regulation;
 KW ISS; cancer; allergy; asthma; hepatitis B infection; papillomavirus;
 KW human immunodeficiency virus; influenza; herpes; M. tuberculosis; ss;
 KW B. pertussis; malaria; plasmodia; Leishmania; Trypanosoma; Schistosoma.
 XX
 OS Synthetic.

XX Key modified_base 11

FT Location/Qualifiers

FT /*tag= a

FT /note= "5-bromocytosine"

XX WO9855495-A2.

PD 10-DEC-1998.

XX 05-JUN-1998; 98WO-US11578.

XX 06-JUN-1997; 97US-0048793.

XX (DYNA-) DYNAVAX TECHNOLOGIES CORP.

XX Dina D, Roman M, Schwartz D;

XX WPI: 1999-059898/05.

PT Immunostimulatory oligonucleotides regulate the immune system - and
 PT contain an immune-stimulating octanucleotide sequence; for treating
 PT cancer, allergic and infectious diseases
 XX
 PS Claim 24; Page 30; 63pp; English.

XX The invention relates to immunomodulatory oligonucleotides that comprise
 CC at least 1 immunostimulatory octanucleotide sequence (ISS) where the ISS
 CC sequences are selected from the group consisting of AACGTTCC, AACGTTCC,
 CC GACGTTCC, and GACGTTCC. The immunomodulatory sequences are used to treat
 CC patients needing immune regulation, such as those suffering from cancer,
 CC an allergic disease and asthma. They are also used to prevent infectious
 CC diseases such as influenza, herpes, hepatitis B, human immunodeficiency
 CC and papillomavirus, Hemophilus influenza, Mycobacterium tuberculosis and
 CC Bordetella pertussis, malarial plasmodia, Leishmania, Trypanosoma and
 CC Schistosoma. The immunomodulatory sequences are used to screen for human
 CC immunostimulatory activity by incubating macrophage cells and the
 CC oligonucleotide; and determining the relative amount of Th1-biased
 CC cytokines in the supernatant. Sequences AAV80096 to AAV80103 represent
 CC specific claimed examples of such immunomodulatory oligonucleotides.

XX Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;
 SQ

Query Match 100.0%; Score 22; DB 20; Length 22;
 Best Local Similarity 100.0%; Pred. No. 0.034;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgactgtgaacgttcgagatga 22
 ||||||||||||||||||
 DB 1 tgactgtgaacgttcgagatga 22

RESULT 6
 AAC64051
 ID AAC64051 standard; DNA; 22 BP.
 XX
 AC AAC64051;
 XX
 DT 15-FEB-2001 (first entry)
 XX
 DE Immunostimulatory Cpg phosphorothioate oligodeoxynucleotide.
 XX
 KW Cpg oligodeoxynucleotide; phosphorothioate; immunostimulatory; ISS ODN;
 KW enhanced antigen presentation; antigen-presenting cell; APC;
 KW T-cell activation; tumour cell; tumour antigen; cancer immunotherapy;
 KW vaccine; ss.
 XX
 OS Synthetic.
 XX
 PN WO200062787-A1.
 XX
 PD 26-OCT-2000.
 XX
 PF 11-APR-2000; 2000WO-US09664.
 XX
 PR 15-APR-1999; 99US-0292278.
 XX
 PA (RESC) UNIV CALIFORNIA.
 XX
 PI Raz E, Martin-Orozco E;
 XX
 DR WPI; 2000-679548/66.
 XX

Enhancing antigen-presentation capabilities of T-cells for cancer immunotherapy, by contacting cells with an immunostimulatory oligonucleotide -

Example 1; Page 18; 42pp; English.

The invention relates to a method of inducing activation of T-cells to respond to an antigen, comprising contacting antigen-presenting cells (APC) with an immunostimulatory oligodeoxynucleotide (ISS-ODN). The APCs thus treated have enhanced antigen presenting capabilities compared to antigen-activated APCs. APCs with enhanced antigen-presentation capabilities then present the antigen to T-cells. The method is useful for cancer immunotherapy. The ISS-ODN is used to enhance the tumour antigen presenting capacity of tumour cells, thereby inducing T-cell activation, and is therefore useful for treating tumours. Additionally, tumour cells treated with an ISS-ODN ex vivo are useful as vaccines. ISS-ODN treated APCs are induced to take up antigen through upregulation of Fe-receptor expression, to present antigen through upregulation of major histocompatibility complex (MHC) Class I and II expression and CD1d expression, to produce co-stimulatory factors (B7 and CD40), to provide cell-to-cell adhesion through upregulation of intercellular adhesion molecule (ICAM) expression, and to increase Th1 stimulatory cytokine production, all at levels greater than that achieved through contact of APC with antigen alone. The present sequence represents a phosphorothioate Cpg ISS-ODN used in the exemplifications of the invention.

Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;

Query Match 100.0%; Score 22; DB 21; Length 22;
 Best Local Similarity 100.0%; Pred. No. 0.034;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgactgtgaacgttcgagatga 22
 ||||||||||||||||||
 DB 1 tgactgtgaacgttcgagatga 22

RESULT 7
 AAA96253
 ID AAA96253 standard; DNA; 22 BP.
 XX
 AC AAA96253;
 XX
 DT 08-FEB-2001 (first entry)
 XX
 DE Sequence of a stabilised oligonucleotide with antitumour activity.
 XX
 KW Antitumour; immunostimulatory oligonucleotide; tumour; anaplasia;
 KW glioblastoma; medullablastoma; neuroblastoma; melanoma; carcinoma; ss.
 XX
 OS Synthetic.
 XX
 PN WO200056342-A2.
 XX
 PD 28-SEP-2000.
 XX
 PF 17-MAR-2000; 2000WO-FR00676.
 XX
 PR 19-MAR-1999; 99FR-0003433.
 XX
 PA (ASSI-) ASSISTANCE PUBLIQUE HOPITAUX PARIS.
 XX
 PI (INRM) INST NAT SANTE & RECH MEDICALE.
 XX
 DR WPI; 2000-602192/57.
 XX

Use of stabilized oligonucleotides as antitumor agents, particularly against nervous system tumors, have optimal activity and are not toxic -

Example 2; Page 16; 57pp; French.

The present sequence represents a stabilised oligonucleotide which has antitumour activity. The oligonucleotide comprises an octamer motif of the type 5'-purine-pyrimidine-Cg-pyrimidine-X-X-3', where the pair X-X is AT, AA, CT or TT. The oligonucleotides are immunostimulatory, and are not toxic. They may be adapted for use in CC animals or humans. The stabilised oligonucleotides are used for treating tumours of any type and any degree of anaplasia, particularly human tumours in the peripheral or central nervous systems, specifically glioblastomas, medullablastomas, neuroblastomas, melanomas or carcinomas.

Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;

Query Match 100.0%; Score 22; DB 21; Length 22;
 Best Local Similarity 100.0%; Pred. No. 0.034;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgactgtgaacgttcgagatga 22
 ||||||||||||||||||
 DB 1 tgactgtgaacgttcgagatga 22

RESULT 8
 AAA90458
 ID AAA90458 standard; DNA; 22 BP.
 XX
 AC AAA90458;

XX 10-JAN-2001 (first entry)
XX Cpg adjutant oligonucleotide, SEQ ID NO:19.
DE
XX
XX Cpg oligonucleotide; Cpg motif; adjuvant; microdroplet emulsion;
KW microemulsion; adsorbent microparticle; vaccine; Th1 immune response;
KW viral infection; bacterial infection; parasitic infection; HCV; HBV;
KW hepatitis C virus; hepatitis B virus; herpes simplex virus; HSV; HTV;
KW human immunodeficiency virus; cytomegalovirus; CMV; influenza virus;
KW rabies virus; cholera; diphtheria; tetanus; pertussis;
KW Helicobacter pylori; Haemophilus influenzae; malaria; ss.
XX
OS Synthetic.
XX
XX WO200050006-A2.
XX
XX 31-AUG-2000.
XX
XX 09-FEB-2000; 2000WO-US03331.
XX
XX 26-FEB-1999; 99US-0121858.
XX 29-JUL-1999; 98US-0146391.
XX 28-OCT-1999; 99US-0161997.
XX
XX (CHIR) CHIRON CORP.
XX
XX O'Hagan D, Ott GS, Donnelly J, Kazaz J, Ugozzoli M, Singh M;
XX Barackman J;
XX WPI: 2000-587123/55.
XX
XX Microemulsion having an adsorbent surface comprising a microdroplet
PT emulsion consisting of a metabolizable oil and an emulsifying agent
PT which is a detergent, useful as a vaccine to treat bacterial, viral,
PT and parasitic infection
XX
XX Claim 17; Page 40; 95pp; English.
XX
XX The invention relates to a microdroplet emulsion (microemulsion) with an
CC adsorbent surface, and which comprises a metabolizable oil and an
CC emulsifying agent (a detergent). It also relates to a composition
CC comprising the microemulsion and a microparticle with an adsorbent
CC surface, where the microparticle comprises a polymer selected from a
CC poly(alpha-hydroxy acid), a polyhydroxy butyric acid, a
CC polycaprolactone, a polyorthoester, a polyhydride, and a
CC polycyanoacrylate, and a second detergent. The surface of the
CC microparticles efficiently adsorb biologically active macromolecules such
CC as DNA, polypeptides, antigens, hormones, pharmaceuticals, enzymes,
CC mediators of transcription or translation, metabolic intermediates and
CC adjuvants. Additionally, a second biologically active molecule may be
CC encapsulated within the microparticle. The microemulsion can be used in
CC methods of immunising a host animal, particularly a human, against a
CC viral, bacterial or parasitic infection, and in methods of increasing a
CC Th1 immune response. The microemulsions (having the appropriate antigens
CC adsorbed) may be particularly used as vaccines for hepatitis C virus
CC (HCV), hepatitis B virus (HBV), herpes simplex virus (HSV), human
CC immunodeficiency virus (HIV), cytomegalovirus (CMV), influenza virus, and
CC rabies virus; the bacteria which cause cholera, diphtheria, tetanus and
CC pertussis; Helicobacter pylori and Haemophilus influenzae; and
CC malaria-causing parasites. Sequences AAAG0447-90467 represent Th1
CC lymphocyte stimulating oligonucleotides containing at least one Cpg motif
CC which are claimed for use as adjuvants in the compositions of the
CC invention.
XX
XX Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;
SQ

Query Match 100.0%; Score 22; DB 21; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.034;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
0y 1 tgactgtgaacgttcgagatga 22

DB 1 tgactgtgaacgttcgagatga 22
XXXXXXXXXXXXXXXXXXXX
RESULT 9
AAAI4467
ID AAAI4467 standard; DNA; 22 BP.
XX
XX AAAI4467;
XX
XX 21-AUG-2000 (first entry)
XX
XX Immunostimulatory oligonucleotide (ISS-ODN) DX1018.
DE
XX
XX Immunostimulatory oligonucleotide; adjuvant; mucosal immunity;
KW secretory immunoglobulin A production; sIgA; Th1 phenotype; ds.
XX
OS Synthetic.
XX
XX WO200020039-A1.
XX
XX 13-APR-2000.
XX
XX 15-SEP-1999; 99WO-US21203.
XX
XX 05-OCT-1998; 98US-0167039.
XX
XX (REGC) UNIV CALIFORNIA.
XX
XX Raz E, Horner AA, Carson DA;
XX
XX WPI: 2000-303647/26.
XX
XX Immunostimulatory oligonucleotide adjuvant induces mucosal immunity to
PT an antigen in a mammalian host through production of secretory
PT immunoglobulin A -
XX
XX
XX Claim 8; Page 21; 64pp; English.
XX
XX The invention relates to a method of inducing mucosal immunity to an
CC antigen in a mammalian host, including the production of secretory
CC immunoglobulin A (sIgA). Immune protection in the mucosa (the principal
CC site of entry of most foreign antigens) is mediated by mucosa-associated
CC lymphoid tissue, epithelial and distinct B-cell, T-cell and accessory
CC cell sub-populations. The primary immune response which characterises
CC the induction of mucosal immunity to an antigen is sIgA production by
CC activated B-cells. The method comprises introducing an immunostimulatory
CC oligonucleotide (ISS-ODN) and the antigen into host mucosa, where the
CC ISS-ODN includes a core nucleotide sequence. The core nucleotide
CC sequence is 5'-Purine-Purine-C-G-Pyrimidine-Pyrimidine-3', specific
CC examples of which are AACGTT, AGCGTC and GACGTT (SEQ ID NOS 1-3). A
CC specific example of an ISS-ODN is DX1018 (AAAI4467). The ISS-ODN is used
CC as an adjuvant with an antigen for stimulating mucosal immunity. The
CC level of sIgA production induced in the host is at least 3 times the
CC magnitude of sIgA production achievable in response to introduction of
CC the antigen alone into the mucosal tissue and is equivalent or greater than
CC the magnitude of sIgA production achievable in response to introduction
CC of the antigen and cholera toxin adjuvant into the mucosal tissue. The
CC host immune response is stimulated to antigen-specific IgA production,
CC biased towards the Th1 phenotype while antigen-induced IgE production is
CC avoided. The adjuvant has little or no known toxicity in mammals and its
CC efficacy is comparable to that of cholera toxin which is used as a
CC mucosal adjuvant. The present sequence represents the immunostimulatory
CC oligonucleotide DX1018.
XX
XX Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;
SQ

Query Match 100.0%; Score 22; DB 21; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.034;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
0y 1 tgactgtgaacgttcgagatga 22

Db 1 tgactgtgaacgttcgagatga 22

RESULT 10
AAA38065
ID AAA38065 standard; DNA: 22 BP.

XX
AC AAA38065;

DT 24-AUG-2000 (first entry)

XX Immunostimulatory sequence (ISS) #1.

DE Immunostimulatory sequence; ISS; immunomodulator; glycoprotein 120;
KW gp120; human immunodeficiency virus; HIV; immune response; infection;
KW development; ss.

XX Synthetic.

OS WO200021556-A1.

PN 20-APR-2000.

PD 08-OCT-1999; 99WO-US23677.

PF 09-OCT-1998; 98US-0103733.

PR 07-OCT-1999; 99US-0415186.

XX (DYNA-) DYNAVAX TECHNOLOGIES CORP.

PI Tighe H, Raz E, Schwartz D, Takabayashi K;

DR WPI; 2000-317846/27.

PT Anti-HIV composition comprises immunostimulatory polynucleotides and
PT HIV glycoprotein gp120 useful for modulating, stimulating an immune
PT response against HIV in an HIV infected individual

XX Claim 3; Page 16; 65pp; English.

XX The present invention relates to an immunostimulatory composition
CC comprising a human immunodeficiency virus (HIV) antigen, and an
CC immunomodulatory polynucleotide comprising an immunostimulatory sequence
CC (ISS). This sequence represents an ISS that can be used in the
CC composition. An immunostimulatory composition which comprises a gp120
CC conjugated to an immunomodulatory polynucleotide, or is proximately
CC associated to it and not conjugated, is used for modulating or
CC stimulating a specific immune response against gp120 in an individual by
CC producing anti-gp120 antibodies or gp120 specific cytotoxic T cells. It
CC is also used for suppressing or delaying development of HIV infection in
CC an individual infected with HIV or an individual at risk of infection in
CC with HIV, respectively. It is also used for treating an individual
CC infected with HIV in need of immune modulation.

XX Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;

Query Match 100.0%; Score 22; DB 21; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.034;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tgactgtgaacgttcgagatga 22
Db 1 tgactgtgaacgttcgagatga 22

RESULT 11

AAA38071
ID AAA38071 standard; DNA: 22 BP.

XX
AC AAA38071;

DT 24-AUG-2000 (first entry)

XX Immunostimulatory sequence (ISS) #7.

DE Immunostimulatory sequence; ISS; immunomodulator; glycoprotein 120;
KW gp120; human immunodeficiency virus; HIV; immune response; infection;
KW development; ss.

XX Synthetic.

OS Key Location/Qualifiers

FT modified_base 11

FT /**tag= a

FT /mod_base= OTHER

FT /note= "5-Bromocytosine"

PN WO200021556-A1.

PD 20-APR-2000.

PF 08-OCT-1999; 99WO-US23677.

PF 09-OCT-1998; 98US-0103733.

PR 07-OCT-1999; 99US-0415186.

XX (DYNA-) DYNAVAX TECHNOLOGIES CORP.

PI Tighe H, Raz E, Schwartz D, Takabayashi K;

DR WPI; 2000-317846/27.

PT Anti-HIV composition comprises immunostimulatory polynucleotides and
PT HIV glycoprotein gp120 useful for modulating, stimulating an immune
PT response against HIV in an HIV infected individual

XX Disclosure; Page 17; 65pp; English.

XX The present invention relates to an immunostimulatory composition
CC comprising a human immunodeficiency virus (HIV) antigen, and an
CC immunomodulatory polynucleotide comprising an immunostimulatory sequence
CC (ISS). This sequence represents an ISS that can be used in the
CC composition. An immunostimulatory composition which comprises a gp120
CC conjugated to an immunomodulatory polynucleotide, or is proximately
CC associated to it and not conjugated, is used for modulating or
CC stimulating a specific immune response against gp120 in an individual by
CC producing anti-gp120 antibodies or gp120 specific cytotoxic T cells. It
CC is also used for suppressing or delaying development of HIV infection in
CC an individual infected with HIV or an individual at risk of infection
CC with HIV, respectively. It is also used for treating an individual
CC infected with HIV in need of immune modulation.

XX Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;

Query Match 100.0%; Score 22; DB 21; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.034;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tgactgtgaacgttcgagatga 22
Db 1 tgactgtgaacgttcgagatga 22

RESULT 12

AAA38072
ID AAA38072 standard; DNA: 22 BP.

XX
AC AAA38072;

DT 24-AUG-2000 (first entry)

XX Immunostimulatory sequence (ISS) #7.

XX		Immunostimulatory sequence; ISS; immunomodulator; glycoprotein 120;
KW	gp120; human immunodeficiency virus; HIV; immune response; infection;	
KM	development; ss.	
XX		
OS	Synthetic.	
XX		
FH	Key	Location/Qualifiers
FT	modified_base	11
FT		/tag= a
FT		/mod_base= OTHER
FT		/note= "5-Bromocytosine"
FT	modified_base	15
FT		/tag= b
FT		/mod_base= OTHER
FT		/note= "5-Bromocytosine"
XX		
PN	WO200021556-A1.	
XX		
PD	20-APR-2000.	
XX		
FF	08-OCT-1999; 99WO-US23677.	
XX		
PR	09-OCT-1998; 98US-0103733.	
PR	07-OCT-1999; 99US-0415186.	
XX		
PA	(DYNA-) DYNAVAX TECHNOLOGIES CORP.	
XX		
PI	Righe H, Raz E, Schwartz D, Takabayashi K;	
XX		
DR	WPI; 2000-317846/27.	
XX		
PT	Anti-HIV composition comprises immunostimulatory polynucleotides and	
PT	HIV glycoprotein gp120 useful for modulating, stimulating an immune	
PT	response against HIV in an HIV infected individual	-
XX		
PS	Disclosure; Page 17; 65pp; English.	
XX		
CC	The present invention relates to an immunostimulatory composition	
CC	comprising a human immunodeficiency virus (HIV) antigen, and an	
CC	immunomodulatory polynucleotide comprising an immunostimulatory sequence	
CC	(ISS). This sequence represents an ISS that can be used in the	
CC	composition. An immunostimulatory composition which comprises a gp120	
CC	conjugated to an immunomodulatory polynucleotide, or is proximately	
CC	associated to it and not conjugated, is used for modulating or	
CC	stimulating a specific immune response against gp120 in an individual by	
CC	producing anti-gp120 antibodies or gp120 specific cytotoxic T cells. It	
CC	is also used for suppressing or delaying development of HIV infection in	
CC	an individual infected with HIV or an individual at risk of infection	
CC	with HIV, respectively. It is also used for treating an individual	
CC	infected with HIV in need of immune modulation.	
XX		
SQ	Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;	
	Query Match	100.0%; Score 22; DB 21; Length 22;
	Best Local Similarity	100.0%; Pred. No. 0.034;
	Matches 22; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
OY	1 tgactgtgaacgttcgagatga 22	
Db	1 tgactgtgaacgttcgagatga 22	
RESULT 13		
AAZ55876		
ID	AAZ55876 standard; DNA; 22 BP.	
XX		
AC	AAZ55876;	
XX		
DT	10-APR-2000 (first entry)	
DE	Immunomodulatory oligonucleotide SEQ ID NO: 1.	
XX		

Query Match	Best Local Similarity	Score 22:	DB 21:	Length 22:
Matches 22: Conservative	100.0%	100.0%	Pred. NO. 0.034;	0; Mismatches 0; Indels 0; Gaps 0
1 tgactgtgaacgttcgagatga 22				
1 tgactgtgaacgttcgagatga 22				

```

XX 01-OCT-2001 (first entry)
DE Phosphorothioate beta-gal/immunostimulatory oligonucleotide.
DE Anaphylactic hypersensitivity; immunomodulatory nucleic acid; vaccine;
KW anaphylaxis-associated symptom; IGE; histamine; phosphorothioate; ss.
XX Synthetic.
OS
XX WO200145750-A1.
PN
XX 28-JUN-2001.
PD
XX 20-DEC-2000; 2000WO-US35064.
PF
XX 21-DEC-1999; 99US-0171830.
PR
XX (REGC ) UNIV CALIFORNIA.
PA
XX Raz E, Horner AA;
PI
XX WPI; 2001-475812/51.
DR
XX Reducing risk of anaphylactic hypersensitivity response to an allergen
PT in a subject, by administering an immunomodulating nucleic acid
PT molecule comprising a specific sequence
XX
XX Example 1; Page 22; 39pp; English.
PS
XX The specification describes a method for reducing a symptom associated
CC with anaphylactic hypersensitivity or risk of anaphylactic response in
CC a subject. The method comprises administering to an individual a
CC nucleic acid molecule comprising an immunomodulatory nucleic acid
CC molecule (INA) comprising the sequence 5'-C-G-3' to reduce
CC anaphylaxis-associated symptom. The method is useful for reducing a
CC symptom associated with anaphylactic hypersensitivity, including a
CC elevated IGE level, elevated histamine level, constriction of the
CC airways and difficult breathing which can lead to anaphylactic reaction
CC or anaphylactic shock, thereby reducing the risk of death. The present
CC sequence represents a beta-gal/immunostimulatory sequence, which was
CC used as a vaccine to protect against the development of anaphylactic
CC hypersensitivity.
XX
SQ Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;

Query Match          100.0%; Score 22; DB 22; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.034;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tgactgtgaacgttcgagatga 22
   |||||||||||||||||||
DB 1 tgactgtgaacgttcgagatga 22

RESULT 15
AAH73439
ID AAH73439 standard; DNA; 22 BP.
XX
AC AAH73439;
XX
XX 01-OCT-2001 (first entry)
DE Immunomodulatory nucleic acid.
XX
XX G3PDH gene; immunomodulatory oligonucleotide; infection; mycobacterium;
KW intracellular pathogen; anti-pathogenic; ss.
XX Unidentified.
XX
XX WO200155341-A2.

```

```

PD 02-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US03029.
PF
XX 31-JAN-2000; 2000US-0179353.
PR
XX (REGC ) UNIV CALIFORNIA.
PA
XX Raz E, Kornbluth R, Catanzaro A, Hayashi T, Carson DA;
PI
XX WPI; 2001-483234/52.
DR
XX
XX Treating infection of intracellular pathogen e.g., Mycobacterium, in a
PT subject, involves administering immunomodulatory nucleic acid molecule
PT to inhibit intracellular replication of intracellular pathogen
XX
XX Examples; Page 26; 54pp; English.
PS
XX The present invention describes a method of treating an infection caused
CC by an intracellular pathogen, involving administering to the patient an
CC immunomodulatory nucleic acid and an anti-pathogenic agent. This is
CC particularly useful in the treatment of mycobacterial infections. The
CC present sequence is an immunomodulatory nucleic acid described in the
CC exemplification of the invention.
XX
SQ Sequence 22 BP; 6 A; 3 C; 7 G; 6 T; 0 other;

Query Match          100.0%; Score 22; DB 22; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.034;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tgactgtgaacgttcgagatga 22
   |||||||||||||||||||
DB 1 tgactgtgaacgttcgagatga 22

```

Search completed: March 11, 2002, 20:59:44
 Job time: 4167 sec

> O <
O I 10 Intelligenetics
> O <

Quest - Quick User-directed Expression Search Tool
Release 5.4

-- Outline of search "8mers_est" --

Selected search type is key against sequence data banks or files.
Selected scope is Sequence.

Selected sequence key from "pappu518.key":

8mers (NA) ID 8mers NA preliminary pattern

1 and-or
2 followed by
3 g or a
3 g or a
3 cg
3 t or c
3 t or c
3 c
3 g or c
2 followed by
3 g or c
3 g or a
3 g or a
3 cg
3 t or c
3 t or c

Selected files:

File : simplex_est.seq

-- Output Parameters --

Format Options: File Options:
Nucleic acid code matching Exact Indirect file
Find non-matching hits only No Sequence or key file
Report key used Yes List of hits
Note position of hit Yes Hit display
Display full annotations Yes Name and annotations
Sequence context 50 Yes

-- Run Parameters --

Run mode Batch
Time to start comparison now
Notify at end of run NO

1 match found in sequence:
bf169293 : TOIG of: bf169293 check: 1397 from: 1 to: 592
(from "simplex_est.seq")
TOIG of: bf169293 check: 1397 from: 1 to: 592

LOCUS BF169293 592 bp mRNA EST 30-OCT-2000
DEFINITION Tm_ad_04B07.SKPL Trichuris muris (parasitic nematode) mixed adult
Trichuris muris cDNA clone Tm_ad_04B07 5' similar to
emb(CAB93501.1) (Y19221) tropomyosin - Anisakis simplex, mRNA
sequence.
ACCESSION BF169293
VERSION BF169293.1 GI:11053904
KEYWORDS EST.
SOURCE Trichuris muris.
ORGANISM Trichuris muris.
Eukaryota; Metazoa; Nematoda; Enoplea; Trichocephalida; Trichuridae
; Trichuris.
REFERENCE 1 (bases 1 to 592)
AUTHORS Blaxter, M.L., Parkinson, J., Whitton, C., Daub, J., Guillian, D., Hall
N., Quayle, M. and Barrell, B.

TITLE Edinburgh University/Sanger Centre Nematode EST project
JOURNAL Unpublished (2000)
COMMENT Contact: Blaxter ML
Institute of Cell, Animal and Population Biology
University of Edinburgh
Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9
3JF, UK.
Tel: +44 131 650 6760
Fax: +44 131 670 5450
Email: mark.blaxter@ed.ac.uk

The library was prepared by Richard Grencis, Manchester University,
Manchester. Sequencing was performed by Claire Whitton, ICAPB,
Edinburgh
PCR Primers
FORWARD: T3
BACKWARD: T7PL
Plate: 04 row: B column: 07
Seq primer: SKPL
High quality sequence stop: 311.
Location/Qualifiers
1. 592
/organism="Trichuris muris"
/db_xref="taxon:70415"
/clone="Tm_ad_04B07"
/clone_1ib="Trichuris muris (parasitic nematode) mixed
adult"
/sex="mixed"
/dev_stage="adult"
/note="Vector: Lambda Zap II; Site_1: EcoRI (5'end);
Site_2: XhoI (3'end); Trichuris muris is a nematode
parasite of rodents related to the human whipworm
Trichuris trichiura. The library was constructed from
Trichuris muris adults (Edinburgh 'E' strain) maintained
in mice, and was provided by Dr. Richard Grencis,
University of Manchester."

BASE COUNT 163 a 141 c 150 g 138 t
ORIGIN

BF169293 length: 592 March 12, 2002 09:13 Type: N Check: 1397
Found using '8mers' (pappu518.key)

167 CTGGTCACGAGAGAGAGATACAGCCATATCCGAGAGCTGATCAGATCCAA |-----|
217 224

227 GAATTGCTGGCTATTATGATGCTCAGTTGAGTAGAGGCGCTT

1 match found in sequence:
bg577835 : TOIG of: bg577835 check: 8229 from: 1 to: 497
(from "simplex_est.seq")
TOIG of: bg577835 check: 8229 from: 1 to: 497

LOCUS BG577835 497 bp mRNA EST 11-APR-2001
DEFINITION Tm_ad_12C11.SKPL Trichuris muris (parasitic nematode) mixed adult
Trichuris muris cDNA clone Tm_ad_12C11 5' similar to
emb(CAB93501.1) (Y19221) tropomyosin - Anisakis simplex, mRNA
sequence.
ACCESSION BG577835
VERSION BG577835.1 GI:1359289
KEYWORDS EST.
SOURCE Trichuris muris.
ORGANISM Trichuris muris.
Eukaryota; Metazoa; Nematoda; Enoplea; Trichocephalida; Trichuridae
; Trichuris.
REFERENCE 1 (bases 1 to 497)
AUTHORS Blaxter, M.L., Parkinson, J., Whitton, C., Daub, J., Guillian, D., Hall
N., Quayle, M. and Barrell, B.
Edinburgh University/Sanger Centre Nematode EST project

JOURNAL
COMMENT

Unpublished (2000)
 Contact: Blaxter ML
 Institute of Cell, Animal and Population Biology
 University of Edinburgh
 Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9
 3J7, UK.
 Tel: +44 131 650 6760
 Fax: +44 131 670 5450
 Email: mark.blaxter@ed.ac.uk
 The library was prepared by Richard Grencis, Manchester University,
 Manchester. Sequencing was performed by the Pathogen Sequencing
 Unit, Sanger Centre, Cambridge, UK (Neil Hall, Mike Quail & Bart
 Barrell).
 PCR Primers
 FORWARD: T3
 BACKWARD: T7PL
 Plate: 12 row: C column: 11
 Seq primer: SKPL
 High quality sequence stop: 486.
 Location/Qualifiers

FEATURES

Source

1. .497
 /organism="Trichoeris muris"
 /db_xref="taxon:70415"
 /clone="Tm.ad.12C11"
 /clone_lib="Trichoeris muris (parasitic nematode) mixed
 adult"
 /sex="mixed"
 /dev_stage="adult"
 /note="Vector: Lambda Zap II; Site.1: EcoRI (5'end);
 Site.2: XhoI (3'end); Trichoeris muris is a nematode
 parasite of rodents related to the human whipworm
 Trichoeris trichiura. The library was constructed from
 Trichoeris muris adults (Edinburgh 'E' strain) maintained
 in mice, and was provided by Dr. Richard Grencis,
 University of Manchester."
 BASE COUNT 123 a 131 c 122 g 121 t
 ORIGIN

BG577835 Length: 497 March 12, 2002 09:13 Type: N Check: 8229 ..
 Found using '8mers' (pappu518.key)

...

102 CTGGTCACGAGAGAGAGACTACAGGCCATATCCGAGAGCTGATCAGACGTTCCAA |-----|
 152 159

162 GAATTGCTGGCTATTATAGTGTCCAGTGGAGTAGAGGCGCTT

...

-- Search Statistics --

Times: CPU Total Elapsed
 00:00:00.00 00:00:00.00
 Number of sequences searched: 24
 Number of sequence hits: 2
 Number of separate matches: 2
 Number of sequence hits saved: 0

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2002, 19:51:12 ; Search time 100.2 Seconds
(without alignments)
49,726 Million cell updates/sec

Title: US-09-802-518-1

Perfect score: 22

Scoring table: 1 tgaactgtgaacgttcgagatga 22

Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA: *
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq: *
3: /cgn2_6/ptodata/1/ina/5A.COMB.seq: *
4: /cgn2_6/ptodata/1/ina/5B.COMB.seq: *
5: /cgn2_6/ptodata/1/ina/5A.COMB.seq: *
6: /cgn2_6/ptodata/1/ina/5B.COMB.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20.4	92.7	22	4	US-09-092-314-2
2	18.8	85.5	22	4	US-09-092-314-1
3	18.8	85.5	22	4	US-09-092-314-3
4	18.8	85.5	22	4	US-09-092-314-10
5	17.2	78.2	22	4	US-09-092-314-4
6	15.6	70.9	22	4	US-09-092-314-5
7	15.6	70.9	22	4	US-09-092-314-7
8	15.6	70.9	22	4	US-09-092-314-8
9	15.6	70.9	1418	1	US-08-391-615-7
10	15.6	70.9	1830	4	US-08-019-931-2
11	15.6	70.9	2505	1	US-08-391-615-1
12	15.6	70.9	6909	2	US-08-804-196-1
13	15.6	70.9	6909	2	US-08-658-340-1
14	15.6	70.9	6909	3	US-08-746-111-26
15	15.2	69.1	1892	2	US-08-933-750C-66
16	15.2	69.1	1892	2	US-08-234-613-66
17	15.2	69.1	6638	2	US-08-070-301-2
18	14.8	67.3	882	1	US-08-622-354-4
19	14.6	66.4	404	4	US-09-060-756-303
20	14.6	66.4	913	2	US-08-975-316-61
21	14.6	66.4	1532	3	US-09-118-324-1
22	14.6	66.4	1614	4	US-09-046-894-29
23	14.6	66.4	2694	4	US-08-975-703-5
24	14.6	66.4	2694	4	US-09-515-884-5
25	14.6	66.4	5408	1	US-08-471-058-20
26	14.6	66.4	5408	3	US-08-471-057-20
27	14.6	66.4	4403765	4	US-09-103-840A-2

28	14.6	66.4	4411529	4	US-09-103-840A-1	Sequence 1, Appl
29	14.2	64.5	864	4	US-08-998-416-297	Sequence 297, App
30	14.2	64.5	1166	1	US-08-121-063-7	Sequence 7, Appl
31	14.2	64.5	176373	3	US-09-128-155-17	Sequence 17, Appl
32	14.2	63.6	77	14	US-08-399-412A-58	Sequence 58, Appl
33	14.2	63.6	95	4	US-08-952-793-258	Sequence 258, App
34	14.2	63.6	95	5	PCT-US96-09455A-258	Sequence 258, App
35	14.2	63.6	657	2	US-08-479-733A-22	Sequence 22, Appl
36	14.2	63.6	657	3	US-08-487-427-22	Sequence 22, Appl
37	14.2	63.6	657	3	US-08-479-727A-22	Sequence 22, Appl
38	14.2	63.6	657	3	US-08-482-369A-22	Sequence 22, Appl
39	14.2	63.6	657	3	US-09-009-217-10	Sequence 10, Appl
40	14.2	63.6	657	3	US-09-009-656-10	Sequence 10, Appl
41	14.2	63.6	657	5	PCT-US95-07439-22	Sequence 22, Appl
42	14.2	63.6	672	1	US-07-816-679A-2	Sequence 2, Appl
43	14.2	63.6	753	5	PCT-US92-11270-2	Sequence 2, Appl
44	14.2	63.6	753	4	US-08-861-774E-73	Sequence 73, Appl
45	14.2	63.6	789	4	US-08-776-059-42	Sequence 42, Appl

ALIGNMENTS

```
RESULT 1
US-09-092-314-2
; Sequence 2, Application US/09092314
; Patent No. 6225292
; GENERAL INFORMATION:
; APPLICANT: Raz, Eyal
; TITLE OF INVENTION: Inhibitors of DNA Immunostimulatory
; Patent No. 6225292
; FILE REFERENCE: 6510-173051
; CURRENT APPLICATION NUMBER: US/09/092.314
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/048.794
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Oligonucleotide
US-09-092-314-2

Query Match: 92.7%; Score 20.4; DB 4; Length 22;
Best Local Similarity 95.5%; Pred. NO. 0.049;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 tgaactgtgaacgttcgagatga 22
Db 1 tgaactgtgaacgttcgagatga 22

RESULT 2
US-09-092-314-1
; Sequence 1, Application US/09092314
; Patent No. 6225292
; GENERAL INFORMATION:
; APPLICANT: Raz, Eyal
; TITLE OF INVENTION: Inhibitors of DNA Immunostimulatory
; Patent No. 6225292
; FILE REFERENCE: 6510-173051
; CURRENT APPLICATION NUMBER: US/09/092.314
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/048.794
; PRIOR FILING DATE: 1997-06-06
```

NUMBER OF SEQ ID NOS: 11
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 22
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Oligonucleotide
US-09-092-314-1

Query Match
Best Local Similarity 85.5%; Score 18.8; DB 4; Length 22;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 tgactgtgaacgttcgagatga 22
|||||
Db 1 tgactgtgaaggttagagatga 22

RESULT 3
US-09-092-314-3
Sequence 3, Application US/09092314
Patent No. 6225292
GENERAL INFORMATION:
APPLICANT: Raz, Eyal
APPLICANT: Roman, Mark
TITLE OF INVENTION: Inhibitors of DNA Immunostimulatory
Patent No. 6225292
FILE REFERENCE: 6510-1730S1
CURRENT APPLICATION NUMBER: US/09/092,314
CURRENT FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/048,794
PRIOR FILING DATE: 1997-06-06
NUMBER OF SEQ ID NOS: 11
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 22
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Oligonucleotide
US-09-092-314-3

Query Match
Best Local Similarity 85.5%; Score 18.8; DB 4; Length 22;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 tgactgtgaacgttcgagatga 22
|||||
Db 1 tgactgtgaaccttagagatga 22

RESULT 4
US-09-092-314-10
Sequence 10, Application US/09092314
Patent No. 6225292
GENERAL INFORMATION:
APPLICANT: Raz, Eyal
APPLICANT: Roman, Mark
TITLE OF INVENTION: Inhibitors of DNA Immunostimulatory
Patent No. 6225292
FILE REFERENCE: 6510-1730S1
CURRENT APPLICATION NUMBER: US/09/092,314
CURRENT FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/048,794
PRIOR FILING DATE: 1997-06-06
NUMBER OF SEQ ID NOS: 11
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10

LENGTH: 22
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Oligonucleotide
US-09-092-314-10

Query Match
Best Local Similarity 85.5%; Score 18.8; DB 4; Length 22;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 tgactgtgaacgttcgagatga 22
|||||
Db 1 tgactgtgaatgttagagatga 22

RESULT 5
US-09-092-314-4
Sequence 4, Application US/09092314
Patent No. 6225292
GENERAL INFORMATION:
APPLICANT: Raz, Eyal
APPLICANT: Roman, Mark
TITLE OF INVENTION: Inhibitors of DNA Immunostimulatory
Patent No. 6225292
FILE REFERENCE: 6510-1730S1
CURRENT APPLICATION NUMBER: US/09/092,314
CURRENT FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/048,794
PRIOR FILING DATE: 1997-06-06
NUMBER OF SEQ ID NOS: 11
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 22
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Oligonucleotide
US-09-092-314-4

Query Match
Best Local Similarity 78.2%; Score 17.2; DB 4; Length 22;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 tgactgtgaacgttcgagatga 22
|||||
Db 1 tgactgtgaagcttagagatga 22

RESULT 6
US-09-092-314-5
Sequence 5, Application US/09092314
Patent No. 6225292
GENERAL INFORMATION:
APPLICANT: Raz, Eyal
APPLICANT: Roman, Mark
TITLE OF INVENTION: Inhibitors of DNA Immunostimulatory
Patent No. 6225292
FILE REFERENCE: 6510-1730S1
CURRENT APPLICATION NUMBER: US/09/092,314
CURRENT FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/048,794
PRIOR FILING DATE: 1997-06-06
NUMBER OF SEQ ID NOS: 11
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5
LENGTH: 22
TYPE: DNA
ORGANISM: Artificial Sequence

FEATURE:
OTHER INFORMATION: Oligonucleotide
US-09-092-314-5

Query Match 70.9%; Score 15.6; DB 4; Length 22;
Best Local Similarity 81.8%; Pred. No. 15;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 tgactgtgaacgttcagagatga 22
|||||
Db 1 tgactgttccttagagatga 22

RESULT 7
US-09-092-314-7
Sequence 7, Application US/09092314
Patent No. 6225292
GENERAL INFORMATION:
APPLICANT: Raz, Eyal
APPLICANT: Roman, Mark
TITLE OF INVENTION: Inhibitors of DNA Immunostimulatory
TITLE OF INVENTION: Sequence Activity
Patent No. 6225292
FILE REFERENCE: 6510-173U51
CURRENT APPLICATION NUMBER: US/09/092,314
CURRENT FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/048,794
PRIOR FILING DATE: 1997-06-06
NUMBER OF SEQ ID NOS: 11
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 22
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Oligonucleotide
US-09-092-314-7

Query Match 70.9%; Score 15.6; DB 4; Length 22;
Best Local Similarity 81.8%; Pred. No. 15;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 tgactgtgaacgttcagagatga 22
|||||
Db 1 tgactgtgaggttcagagatga 22

RESULT 8
US-09-092-314-8
Sequence 8, Application US/09092314
Patent No. 6225292
GENERAL INFORMATION:
APPLICANT: Raz, Eyal
APPLICANT: Roman, Mark
TITLE OF INVENTION: Inhibitors of DNA Immunostimulatory
TITLE OF INVENTION: Sequence Activity
Patent No. 6225292
FILE REFERENCE: 6510-173U51
CURRENT APPLICATION NUMBER: US/09/092,314
CURRENT FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/048,794
PRIOR FILING DATE: 1997-06-06
NUMBER OF SEQ ID NOS: 11
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 8
LENGTH: 22
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Oligonucleotide
US-09-092-314-8

Query Match 70.9%; Score 15.6; DB 4; Length 22;
Best Local Similarity 81.8%; Pred. No. 15;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 tgactgtgaacgttcagagatga 22
|||||
Db 1 tgactgtgaggttcagagatga 22

RESULT 9
US-08-391-615-7
Sequence 7, Application US/08391615
Patent No. 5550054
GENERAL INFORMATION:
APPLICANT: Witte, Owen
APPLICANT: Tsukada, Satoshi
APPLICANT: Saffran, Douglas
APPLICANT: Rawlings, David
TITLE OF INVENTION: HEMATOPOIETIC RESTRICTED TYROSINE KINASE
TITLE OF INVENTION:
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/391,615
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/006,449
FILING DATE: 21-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Rowland, Bettam I
REGISTRATION NUMBER: 20,015
REFERENCE/DOCKET NUMBER: A-57583-1/BIR UCLA 232-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277289 FHT UR
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1418 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-391-615-7

Query Match 70.9%; Score 15.6; DB 1; Length 1418;
Best Local Similarity 81.8%; Pred. No. 27;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 tgactgtgaacgttcagagatga 22
|||||
Db 218 TGACTTGAACGTGGAGAGA 239

RESULT 10
US-09-019-931-2
Sequence 2, Application US/09019931
Patent No. 6194148

GENERAL INFORMATION:
APPLICANT: Hori, Kunio, Takahashi, Takeo, Okada, Takao
TITLE OF INVENTION: A Method For Detecting A Hybridized
TITLE OF INVENTION: Nucleic Acid Molecule
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fishauf, Holtz, Goodman, Langer & Chick
STREET: 767 Third Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10017-2023
COMPUTER READABLE FORM:
MEDIUM TYPE: 3+ inch, 1.44 mb diskette
COMPUTER: IBM PC
OPERATING SYSTEM: MS DOS
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/019,931
FILING DATE: 06-FEB-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 9-025291
FILING DATE: 07-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Barth, Richard
REGISTRATION NUMBER: 28,180
REFERENCE/DOCKET NUMBER: 980048/HG
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 319-4900
TELEFAX: (212) 319-5101
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1830
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
IMMEDIATE SOURCE:
LIBRARY: PGEX-PH
US-09-019-931-2

Query Match 70.9%; Score 15.6; DB 4; Length 1830;
Best Local Similarity 81.8%; Pred. No. 28;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 tgactgtgaacgttcgagatga 22
||||| ||||| ||||| ||
Db 881 TGACTTGAACGTGCGAGAGA 902

RESULT 11
US-08-391-615-1
Sequence 1, Application US/08391615
Patent No. 5550054
GENERAL INFORMATION:
APPLICANT: Witte, Owen
APPLICANT: Tsukada, Satoshi
APPLICANT: Saffran, Douglas
APPLICANT: Rawlings, David
TITLE OF INVENTION: HEMATOPOIETIC RESTRICTED TYROSINE KINASE
TITLE OF INVENTION:
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/391,615
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/006,449
FILING DATE: 21-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Rowland, Bertlam I
REGISTRATION NUMBER: 20,015
REFERENCE/DOCKET NUMBER: A-57583-1/BIR UCIA 232-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299 FHT UR
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2505 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 137..2116
US-08-391-615-1

Query Match 70.9%; Score 15.6; DB 1; Length 2505;
Best Local Similarity 81.8%; Pred. No. 29;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 tgactgtgaacgttcgagatga 22
||||| ||||| ||||| ||
Db 262 TGACTTGAACGTGCGAGAGA 283

RESULT 12
US-08-804-196-1/c
Sequence 1, Application US/08804196
Patent No. 5874256
GENERAL INFORMATION:
APPLICANT: Bertina, Rogier
APPLICANT: Reitsma, Pletier
TITLE OF INVENTION: A method for diagnosing an increased
TITLE OF INVENTION: risk for thrombosis or a genetic defect causing
TITLE OF INVENTION: thrombosis and kit for use with the same.
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Akzo No. 5874256e1 Patent Department
STREET: 1300 Piccard Drive, Suite 206
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804,196
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/454,353
FILING DATE: 06-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Gormley, Mary E.
REGISTRATION NUMBER: 34,409
TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 948-7400
TELEFAX: (301) 948-9751
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6909 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: unknown
MOLECULE TYPE: CDNA
US-08-804-196-1

Query Match 70.9%; Score 15.6; DB 2; Length 6909;
Best Local Similarity 81.8%; Pred. No. 33;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 tgactgtgacgttcgagatga 22
|||||
DB 793 TGACTGTGTACATTAGGATGA 772

RESULT 13
US-08-658-340-1/c
Sequence 1, Application us/08658340
Patent No. 5910576
Patent No. 5910576 5861489

GENERAL INFORMATION:
APPLICANT: Bertina, Rogier
APPLICANT: Reitsma, Pieter
TITLE OF INVENTION: A method for diagnosing an increased
TITLE OF INVENTION: risk for thrombosis or a genetic defect causing
TITLE OF INVENTION: thrombosis and kit for use with the same.
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESS: Akzo No. 5910576 5861489el Patent Department
STREET: 1300 Piccard Drive, Suite 206
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/658,340
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/454,353
FILING DATE: 06-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Gormley, Mary E.
REGISTRATION NUMBER: 34,409
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 948-7400
TELEFAX: (301) 948-9751
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6909 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: unknown
MOLECULE TYPE: CDNA
US-08-658-340-1

Query Match 70.9%; Score 15.6; DB 2; Length 6909;
Best Local Similarity 81.8%; Pred. No. 33;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 tgactgtgacgttcgagatga 22
|||||

DB 793 TGACTGTGTACATTAGGATGA 772

RESULT 14
US-08-746-111-26/c
Sequence 26, Application US/08746111
Patent No. 6066778

GENERAL INFORMATION:
APPLICANT: Ginsburg, David
APPLICANT: Cul, Jisong
TITLE OF INVENTION: Compositions And Methods For Screening
TITLE OF INVENTION: Compounds For Anticoagulant Activity
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible-
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/746,111
FILING DATE: 06-NOV-1996
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: UM-02536

TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 6909 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-746-111-26

Query Match 70.9%; Score 15.6; DB 3; Length 6909;
Best Local Similarity 81.8%; Pred. No. 33;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0

QY 1 tgactgtgacgttcgagatga 22
|||||
DB 793 TGACTGTGTACATTAGGATGA 772

RESULT 15
US-08-933-750C-66/c
Sequence 66, Application US/08933750C
Patent No. 5932442

GENERAL INFORMATION:
APPLICANT: Lal, Preeti
APPLICANT: Hillman, Jennifer L.
APPLICANT: Bandman, Olga
APPLICANT: Shah, Purvi
APPLICANT: Au-Young, Janice
APPLICANT: Yue, Henry
APPLICANT: Guegler, Karl J.
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
NUMBER OF SEQUENCES: 98
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive

CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94304
 COMPUTER READABLE FORM
 MEDIUM TYPE: Disk
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/933,750C
 FILING DATE: September 23, 1997
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Billings, Lucy J.
 REGISTRATION NUMBER: 36,749
 REFERENCE/DOCKET NUMBER: PF-0356 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-855-0555
 TELEFAX: 415-845-4166
 TELEX:
 INFORMATION FOR SEQ ID NO: 66:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1892 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 LIBRARY: BRSTN0703
 CLONE: 641127
 US-08-933-750C-66

Query Match 69.1%; Score 15.2; DB 2; Length 1892;
 Best Local Similarity 85.0%; Pred. No. 45;
 Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 tgactgtgaagcttcagat 20
 ||| ||||| ||||| ||
 DB 1278 TGAATGTGAAGCTTCAGCT 1259

Search completed: March 11, 2002, 21:01:45
 Job time: 4233 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2002, 19:41:02 ; Search time 2230.08 Seconds

(without alignments)
106.008 Million cell updates/sec

Title: US-09-802-518-1

Perfect score: 22

Sequence: 1 tgaactgtgaagcttcgagatga 22

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

EST.*
1: em_estfun:*
2: em_esthum:*
3: em_estlin:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_hlc:*
10: gb_estl:*
11: gb_est2:*
12: gb_hlc:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rtd:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	17.8	80.9	561	13	AZ755668
2	17.8	80.9	961	11	BF971856
3	17.4	79.1	489	13	AZ060178
4	17.4	79.1	530	13	AZ886419
5	17.2	78.2	374	13	AQ245026
6	17.2	78.2	408	13	AZ536502
7	17.2	78.2	424	10	BE723539
8	17.2	78.2	463	10	AU083559
9	17.2	78.2	479	10	AU089685
10	17.2	78.2	972	13	CNS05PD9
11	16.8	76.4	105	10	AA094019
12	16.8	76.4	523	13	AZ483488

13	16.8	76.4	526	13	AZ501799	AZ501799
14	16.8	76.4	681	10	AV732648	AV732648
15	16.8	76.4	705	10	AW16461	AW16461
16	16.4	74.5	400	10	AW398307	AW398307
17	16.4	74.5	496	10	AW034934	AW034934
18	16.4	74.5	546	13	AZ058706	AZ058706
19	16.4	74.5	554	13	AZ280611	AZ280611
20	16.4	74.5	559	13	AZ068022	AZ068022
21	16.4	74.5	600	13	AZ976014	AZ976014
22	16.4	74.5	633	11	BG570577	BG570577
23	16.4	74.5	747	11	BG127461	BG127461
24	16.4	74.5	924	13	CNS027SC	CNS027SC
25	16.2	73.6	202	10	AA236074	AA236074
26	16.2	73.6	226	10	BB183285	BB183285
27	16.2	73.6	236	11	BF932252	BF932252
28	16.2	73.6	251	10	BB565758	BB565758
29	16.2	73.6	286	10	AI099019	AI099019
30	16.2	73.6	297	10	AA445764	AA445764
31	16.2	73.6	300	11	CI1370	CI1370
32	16.2	73.6	304	13	AQ581066	AQ581066
33	16.2	73.6	317	10	BE119339	BE119339
34	16.2	73.6	360	10	AV189436	AV189436
35	16.2	73.6	398	10	BE428418	BE428418
36	16.2	73.6	401	10	BE275964	BE275964
37	16.2	73.6	416	13	AO646593	AO646593
38	16.2	73.6	422	11	BG813145	BG813145
39	16.2	73.6	434	11	BF851024	BF851024
40	16.2	73.6	441	10	AI597068	AI597068
41	16.2	73.6	448	10	AW390277	AW390277
42	16.2	73.6	452	11	BG814270	BG814270
43	16.2	73.6	457	10	BE367478	BE367478
44	16.2	73.6	472	13	AQ904593	AQ904593
45	16.2	73.6	473	11	R83554	R83554

ALIGNMENTS

RESULT 1
AZ755668/c
LOCUS
DEFINITION
ev02g09.x1 PAX3 CASTING Library 'ev' Homo sapiens genomic clone
ACCESSION
AZ755668
VERSION
KEYWORDS
GSS.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 561)
Barber,T.D., Barber,M.C., Tomescu,O., Barr,F., Ruben,S. and Friedman,T.B.
TITLE
Cyclic amplification and selection of target genes regulated by Pax3 and PAX3/RRR in embryogenesis and alveolar rhabdomyosarcoma
JOURNAL
Unpublished (2000)
COMMENT
Contact: Friedman TB
Laboratory of Molecular Genetics
National Institute on Deafness and Other Communication Disorders,
National Institutes of Health
5 Research Court, Room 2A-15, Rockville, MD 20850, USA
Tel: 301 402 7580
Fax: 301 496 7882
Email: friedman@nidcd.nih.gov
Plate: 02 row: 9 column: 09
Seq primer: -21M13 forward primer (ABI)
Class: random plasmid subclone.
Location/Qualifiers
1..561
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="ev02g09"
/clone_lib="PAX3 CASTING Library 'ev'"

/sex="Male"
/lab_host="DH10B"
/note="Vector: pGEM-T Easy; Human genomic DNA was partially digested with Sau3AI, ligated to ds linkers, and enriched for binding to human PAX3d0+ protein using a whole genome PCR-based strategy. DNA fragments containing putative PAX3d0+ binding sites were amplified by PCR and cloned into pGEM-T Easy (Promega). The ligation products were transformed into DH10B electrocompetent cells (Life Technologies)."

BASE COUNT 126 a 149 c 152 g 134 t
ORIGIN

Query Match 80.9%; Score 17.8; DB 13; Length 561;
Best Local Similarity 90.5%; Pred. No. 1.8e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 tgactgtgaacttcgagatg 21
|||||
Db 461 TGACTGTGAACTTCAGAGTGA 441

RESULT 2
BF971856 961 bp mRNA EST 22-JAN-2001
LOCUS 602240444F1 NIH_MGC_46 Homo sapiens CDNA IMAGE:4328890 5',
DEFINITION mRNA sequence.
ACCESSION BF971856
VERSION BF971856.1 GI:12339071
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 961)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-femail.nih.gov
Tissue Procurement: ATCC

CNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LMNL at:
http://image.lnl.gov
Plate: L10CM1189 row: h column: 11
High quality sequence stop: 555.

FEATURES

Location/Qualifiers
1..961
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="4328890"
/clone_id="NIH_MGC_46"
/tissue_type="leiomysarcoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: uterus; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

BASE COUNT 205 a 225 c 240 g 291 t
ORIGIN

Query Match 80.9%; Score 17.8; DB 11; Length 961;
Best Local Similarity 90.5%; Pred. No. 2.1e+02;

Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 2 gactgtgaacttcgagatga 22
|||||
Db 650 GACTGTGAACCTCCGATGA 670

RESULT 3
A2060178 489 bp DNA GSS 30-MAR-2000
LOCUS RPCI-23-405E23.TU RPCI-23 Mus musculus genomic clone RPCI-23-405E23
DEFINITION 'DNA sequence.
ACCESSION A2060178
VERSION A2060178.1 GI:7351427
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 489)
AUTHORS Zhao, S., Nierman, W., Feldblum, T., Malek, J., Shatsman, S., Akiret, B., Levins, M., McGann, S., Tesgaye, G., Geer, K., Krol, M., de Jong, P., and Fraser, C. M.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

COMMENT Mouse BAC End Sequences from Library RPCI-23
Unpublished (1999)
Other_GSSs: RPCI-23-405E23.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208

Email: szhao@tigr.org
Emails are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong (pieter@edj.org.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/orderingfrmme.htm) or from Resea ch Genetics (info@resgen.com). BAC end page: http://www.tigr.org/tdb/bac/ends/mouse/bac_end_intro.html
Plate: 405 row: E column: 23
Seq primer: SP6
Class: BAC ends.

FEATURES

Location/Qualifiers
1..489
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_image="RPCI-23-405E23"
/clone_id="RPCI-23"
/sex="Female"
/lab_host="DH10B"
/note="Organ: Kidney/Brain; Vector: pBACe3 6; Site_1: EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI methylase. Size selected DNA was cloned into the pBACe3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."

BASE COUNT 162 a 61 c 67 g 198 t
ORIGIN

Query Match 79.1%; Score 17.4; DB 13; Length 489;
Best Local Similarity 94.7%; Pred. No. 2.7e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 tgactgtgaacttcgaga 19
|||||
Db 170 TGACTGTGAACATTCGACA 152

RESULT 4
A2886419/c

LOCUS A2886419 530 bp DNA GSS 05-MAR-2001
 DEFINITION RPCI-23-18216.TV RPCI-23 Mus musculus genomic clone RPCI-23-18216,
 DNA sequence.
 ACCESSION A2886419
 VERSION A2886419.1 GI:13205364
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 530)
 Zhao, S., Niernman, W., Feldlyum, T., Malek, J., Shatsman, S., Akiret
 'B., Levins, M., McGann, S., Isegaye, G., Geer, K., Krol, M., de Jong, P.
 and Fraser, C.M.
 Mouse BAC End Sequences from Library RPCI-23
 Unpublished (1999)
 Other GSSs: RPCI-23-18216.TV
 Contact: Shaying Zhao
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: szhao@tigr.org
 Clones are derived from the mouse BAC library RPCI-23. For BAC
 library availability, please contact Pieter de Jong
 (pdejong@mail.cho.org). Clones may be purchased from BACPAC
 Resources (<http://www.choi.org/bacpac/orderingframe.html>). BAC end
 page: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
 Plate: 182 row: 1 column: 6
 Seq primer: SP6
 Class: BAC ends.
 FEATURES
 source Location/Qualifiers
 1..530
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="RPCI-23-18216"
 /clone_1lb="RPCI-23"
 /sex="Female"
 /lab_host="DH10B"
 /note="Organ: Kidney/Brain; Vector: pBAC3.6; Site: 1;
 EcoRI, Site: 2; EcoRI; Female C57BL/6J mouse kidney and/or
 brain genomic DNA was isolated and partially digested
 with a combination of EcoRI and EcoRI Methylase. Size
 selected DNA was cloned into the pBAC3.6 vector at the
 EcoRI sites. The ligation products were transformed into
 DH10B electrocompetent cells (BRL Life Technologies)."

BASE COUNT 174 a 65 c 73 g 218 t
 ORIGIN

Query Match 79.1%; Score 17.4; DB 13; Length 530;
 Best Local Similarity 94.7%; Pred. No. 2.8e+02;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Oy 1 tgactgtgaacgttcgaga 19
 ||||||||| |||||
 Db 189 TGACTGTGACATTCGAGA 171

RESULT 5
 A0245026 374 bp DNA GSS 03-OCT-1998
 LOCUS HS_2056_B1_E03_MR_C1T Approved Human Genomic Sperm Library D Homo
 DEFINITION sapiens genomic clone Plate-2056 Col-5 Row-J, DNA sequence.
 ACCESSION A0245026
 VERSION A0245026.1 GI:3691600
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 374)
 AUTHORS Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,
 Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and
 Hood, L.
 TITLE Sequence-tagged connectors: A sequence approach to mapping and
 scanning the human genome
 JOURNAL Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
 MEDLINE 99380589
 COMMENT Contact: Mahairas GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu
 Sequence Tagged Connector
 Plate: 2056 row: J column: 5
 Class: BAC ends
 High quality sequence stop: 374.
 FEATURES
 source Location/Qualifiers
 1..374
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="Plate-2056 Col-5 Row-J"
 /clone_1lb="C1T Approved Human Genomic Sperm Library D"
 /sex="male"
 /note="Organ: sperm; Vector: pBelOBAC11; BAC Clones in
 E-Coli DH10B"

BASE COUNT 91 a 73 c 92 g 117 t 1 others
 ORIGIN

Query Match 78.2%; Score 17.2; DB 13; Length 374;
 Best Local Similarity 86.4%; Pred. No. 3.2e+02;
 Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Oy 1 tgactgtgaacgttcgagatga 22
 ||||||||| |||||
 Db 207 TGACTGTGACGATTCGATCA 228

RESULT 6
 A2536502 408 bp DNA GSS 03-NOV-2000
 LOCUS A2536502
 DEFINITION 110300_96 Planococcus lillacinus DNA Planococcus lillacinus genomic,
 DNA sequence.
 ACCESSION A2536502
 VERSION A2536502.1 GI:11093449
 KEYWORDS GSS.
 SOURCE lillac mealbug.
 ORGANISM Planococcus lillacinus
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha;
 Aphidiformes; Coccoidea; Pseudococcidae; Planococcus.
 1 (bases 1 to 408)
 Mohan, K.N. and Chandra, B.S.
 Mealbug shotgun sequencing
 Unpublished (2000)
 Contact: Mohan KN
 Microbiology and Cell Biology
 Indian Institute of Science
 Sir C.V. Raman Avenue, Bangalore, Karnataka 560012, India
 Email: mohan@cmb1.iisc.ernet.in
 Class: shotgun.
 FEATURES
 source Location/Qualifiers
 1..408
 /organism="Planococcus lillacinus"
 /db_xref="taxon:40930"
 /clone_1lb="Planococcus lillacinus DNA"

BASE COUNT 134 a 83 c 80 g 111 t
 ORIGIN


```

/clone.lib="Rice callus"
/Note="Vector: pBluescript II SK+; Site.1: SalI; Site.2:
NotI; cDNA prepared from rice callus mRNAs by using
oligo(dT) as a primer and ligated to the SalI-NotI site
of pBluescript II SK+ phagemid."
BASE COUNT      149 a      96 c      104 g      130 t
ORIGIN

Query Match.      78.2%; Score 17.2; DB 10; Length 479;
Best Local Similarity 86.4%; Pred. No. 3.4e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1  tgactgtgaactgcagatga 22
        ||| ||||| ||| ||||| |||
DB      299  TGAGTGTGAATGTTAGAGATGA 320

RESULT 10
CNS05PD9/c      972 bp      DNA      GSS      26-MAY-2000
LOCUS      Tetraodon nigroviridis genome survey sequence T7 end of clone
DEFINITION      005F08 of library A from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION      AL347814.1 GI:8241584
VERSION      GSS; genome survey sequence.
KEYWORDS      Tetraodon nigroviridis.
SOURCE      Tetraodon nigroviridis.
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Acanthopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
REFERENCE      1 (bases 1 to 972)
AUTHORS      Roest-Crolius,H., Jallion,O., Dasilva,C., Fizames,C., Fisher,C.,
Bonneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
Weissenbach,J.
TITLE      Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
JOURNAL      Unpublished
REFERENCE      2 (bases 1 to 972)
AUTHORS      Roest-Crolius,H., Jallion,O., Dasilva,C., Fizames,C., Fisher,C.,
Bernot,A., Fizames,C., Wincker,P., Brothier,P., Quetier,F.,
Saurin,W. and Weissenbach,J.
TITLE      Human gene number estimate provided by genome wide analysis using
Tetraodon nigroviridis DNA sequence
JOURNAL      Unpublished
REFERENCE      3 (bases 1 to 972)
AUTHORS      Genoscope.
TITLE      Direct Submission
COMMENT      Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetraodon.
FEATURES
Source
1..972
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone="005F08"
/clone.lib="A"
/Note="Genoscope sequence ID : CONA005DC04C1-end : T7"
BASE COUNT      195 a      268 c      199 g      297 t      13 others
ORIGIN

Query Match.      78.2%; Score 17.2; DB 13; Length 972;
Best Local Similarity 86.4%; Pred. No. 4.1e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1  tgactgtgaactgcagatga 22
        ||| ||||| ||| ||||| |||
DB      46  TGCGTGTGAAGTGCAGATGA 25

```

```

RESULT 11
AA094019      105 bp      mRNA      EST      25-OCT-1996
LOCUS      c11619.seq.F Human fetal heart, Lambda ZAP Express Homo sapiens
DEFINITION      CDNA 5', mRNA sequence.
ACCESSION      AA094019
VERSION      AA094019.1 GI:1639612
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE      1 (bases 1 to 105)
AUTHORS      Llew,C.C.
TITLE      cDNAs from fetal heart (1996)
JOURNAL      Unpublished (1996)
COMMENT      Contact: Llew CC
Department of Laboratory Medicine and Pathobiology
University of Toronto
Bartling Institute, 100 College St., Toronto, Ontario, M5G1L5
Tel: 4169788758
Fax: 4169785650
Email: llew@celltoc.utoronto.ca
PCR PRIMERS
FORWARD: 5' GCCAAGCTCGAATTAACCTCAGCTAAAGG 3'
BACKWARD: 5' CCAGGATGTGTAATGACGACCTCACTATAGGCG 3'
Seq primer: 5' GAATTAACCTCAGCTAAAGG 3'
FEATURES
Source
1..105
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone.lib="Human fetal heart, Lambda ZAP Express"
/lab_host="E. coli XL1-BLue"
/Note="Vector: Lambda ZAP Express; Site.1: EcoRI; Site.2:
XhoI; mRNA was purified from human fetal hearts (8-10
weeks). cDNA was synthesized using a XhoI-Oligo dT
adaptor-primer. EcoRI adaptors were ligated, followed by
digestion with XhoI, for directional cloning into
predigested lambda ZAP Express."
BASE COUNT      31 a      21 c      16 g      37 t
ORIGIN

Query Match.      76.4%; Score 16.8; DB 10; Length 105;
Best Local Similarity 90.0%; Pred. No. 3.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1  tgactgtgaactgcagat 20
        ||| ||||| ||| ||||| |||
DB      43  TGACTGTGAACCTTCAGAT 62

RESULT 12
AA483488      523 bp      DNA      GSS      05-OCT-2000
LOCUS      IM0309M12F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
DEFINITION      clone UUGC1M0309M12 F, DNA sequence.
ACCESSION      AA483488
VERSION      AA483488.1 GI:10647510
KEYWORDS      GSS.
SOURCE      house mouse.
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 523)
AUTHORS      Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
TITLE      Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

```

JOURNAL
COMMENT

Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0309 row: M column: 12
Seq primer: CCTGTAAACACGACGCCAGT
Class: plasmid ends
High quality sequence stop: 523.

FEATURES

source

1. 523
Location/Qualifiers
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0309M12"
/clone_1lb="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g114732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT
ORIGIN

162 a 127 c 103 g 131 t

Query Match

76.4%; Score 16.8; DB 13; Length 523;

Best Local Similarity 90.0%; Pred. No. 5.4e+02;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 actgtgaacgttcgagatga 22
||||| |||||||
Db 376 ACTGTGACTTTCGAGATGA 395

RESULT 13

AZ501799 526 bp DNA GSS 05-OCT-2000
LOCUS 1M0340J17R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
DEFINITION clone UUGC1M0340J17 R, DNA sequence.
ACCESSION AZ501799
VERSION AZ501799.1 GI:10683115
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS

Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
and Wright, D., Weiss, R.
Mouse whole genome scaffolding with paired end reads from 10kb

JOURNAL
COMMENT

Plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0340 row: J column: 17
Seq primer: CACACAGGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 526.

FEATURES

source

1. 526
Location/Qualifiers
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0340J17"
/clone_1lb="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g114732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT
ORIGIN

121 a 126 c 139 g 140 t

Query Match

76.4%; Score 16.8; DB 13; Length 526;

Best Local Similarity 90.0%; Pred. No. 5.4e+02;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 tgactgtgaacgttcgagat 20
||||| |||||||
Db 60 TGACAGTGACGTTCTAGAT 79

RESULT 14

AV732648 681 bp mRNA EST 17-OCT-2000
LOCUS AV732648 HTF Homo sapiens cDNA clone HTFBD03 5', mRNA sequence.
DEFINITION AV732648
ACCESSION AV732648
VERSION AV732648.1 GI:10650193
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS

Gu, Y., Peng, Y., Song, H., Huang, Q., Yang, Y., Gao, G., Xiao, H., Xu, X.,
Li, N., Qian, B., Liu, F., Qu, D., Gao, X., Cheng, Z., Xu, Z., Zeng, L., Xu,
S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G., Hu, R.,
Chen, J., Chen, Z. and Han, Z.
Homo sapiens cDNA HTF clones

JOURNAL
COMMENT

Unpublished (2000)
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzgchgc.sh.cn

FEATURES

source

This clone is available at CHGC in Shanghai.

Location/Qualifiers

1..661
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HTF"
/tissue_type="Hypothalamus"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT 202 a 137 c 156 g 184 t 2 others
ORIGIN

Query Match 76.4%; Score 16.8; DB 10; Length 681;
Best Local Similarity 90.0%; Pred. No. 5.8e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 3 actgtgaacgttcgagatga 22
||||| || |||||
DB 382 ACTGTGAACACTTCGAGATGA 363

RESULT 15
AW916461

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AW916461 705 bp mRNA EST 25-MAY-2000
EST347765 Rat gene index, normalized rat, norvegicus, Bento Soares
Rattus norvegicus cdna clone RGID049 5' end, mRNA sequence.
AW916461
AW916461.1 GI:8082187
EST.
Norway rat.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 705)
Lee, N.H., Glodek, A., Chandra, I., Mason, T.M., Quackenbush, J.,
Kerlavage, A.R. and Adams, M.D.
Rat Genome Project: Generation of a Rat EST (REEST) Catalog & Rat
Gene Index
Unpublished (1998)
Contact: Lee, NH
The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-3529
Fax: (301)-838-0208
Email: nhlee@tigr.org
This clone is available through the ATCC, contact the ATCC
tel#703-365-2700 for further information
Seq primer: M13 Reverse

FEATURES

source

Location/Qualifiers
1..705
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone_lib="RGID049"
/clone_lib="Rat gene index, normalized rat, norvegicus,
Bento Soares"
/tissue_type="mix - brain, ovary, placenta, kidney, lung,
liver, embryo, heart, muscle, spleen"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; Estimated insert size approx.1 kb"

BASE COUNT 177 a 201 c 169 g 157 t 1 others
ORIGIN

Query Match 76.4%; Score 16.8; DB 10; Length 705;
Best Local Similarity 90.0%; Pred. No. 5.9e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 3 actgtgaacgttcgagatga 22
||||| || |||||
DB 150 ACTGTGACCTTCGAGATGA 169

Search completed: March 11, 2002, 20:30:00
JOB time: 2938 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2002, 20:30:00 : Search time 2230.08 Seconds
(without alignments)
115.645 Million cell updates/sec

Title: US-09-802-518-9

Perfect score: 24
Sequence: 1 tcgtcgacgttcgttaacgtcg 24

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: em_estfun:*
2: em_esthum:*
3: em_estlin:*
4: em_estom:*
5: em_estopl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_hic:*
10: gb_estl:*
11: gb_est2:*
12: gb_hic:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_fod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	17.6	73.3	531	10	AV403813 AV403813
2	17.6	73.3	571	10	AL118705 DKFZP761L
3	17.6	73.3	609	13	AO623639 HS.5377.A
4	17.2	71.7	794	13	CNS02XJ3
5	17.2	71.7	2228	11	BR860303
6	16.6	69.2	360	11	D37092
7	16.6	69.2	379	11	C12646
8	16.6	69.2	460	10	AI531538
9	16.6	69.2	464	13	AO860867
10	16.6	69.2	475	10	AW942684
11	16.6	69.2	522	10	AA517098
12	16.6	69.2	532	11	B638934

13	16.6	69.2	537	10	AI976120
14	16.6	69.2	552	11	BG795168
15	16.6	69.2	576	11	BG799488
16	16.6	69.2	586	11	BG799472
17	16.6	69.2	598	11	BF815578
18	16.6	69.2	672	10	AW017352
19	16.6	69.2	696	11	BG641369
20	16.6	69.2	713	13	CNS03K6W
21	16.6	69.2	799	11	BG638867
22	16.6	69.2	893	10	BE413349
23	16.6	69.2	947	11	BG823392
24	16.6	69.2	1095	11	BF686639
25	16.6	69.2	1892	11	BG847877
26	16.4	68.3	454	11	BG639770
27	16.4	68.3	1019	11	BF682153
28	16.2	67.5	546	10	AI437242
29	16.2	67.5	548	13	AO968186
30	16.2	67.5	611	11	BG891803
31	16.2	67.5	624	13	AO968185
32	16.2	67.5	639	11	BG304106
33	16.2	67.5	644	10	AM154870
34	16.2	67.5	644	11	BF677194
35	16.2	67.5	701	13	AO362660
36	16.2	67.5	870	13	CNS0118T
37	16.2	67.5	913	13	CNS01V04
38	16.2	67.5	987	11	BG176913
39	16.2	67.5	1051	13	CNS04PA3
40	16.2	67.5	1083	13	CNS051V8
41	16	66.7	360	11	D67466
42	16	66.7	368	10	AU036374
43	16	66.7	384	10	AA123623
44	16	66.7	399	13	AO852283
45	16	66.7	457	13	TA364A06P

ALIGNMENTS

RESULT 1
AV403813/C
LOCUS
DEFINITION
AV403813 Bombyx mori pheromone gland Shuko x Ryuhaku newly-released
adult Bombyx mori cDNA clone pg--0086 T3, mRNA sequence.
AV403813
AV403813.1 GI:6907901

ACCESSION
AV403813
VERSION
KEYWORDS
SOURCE
ORGANISM
domestic silkworm.
Bombyx mori

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Mita, K., Morimoto, M., Shimada, T., Okano, K. and Maeda, S.
Bombyx mori cDNA
Unpublished (2000)
Contact: Mita K
Genome Research Group
National Institute of Radiological Sciences
Anagawa 4-9-1, Inage, Chiba 263-8555, Japan
Email: kmita@nirs.go.jp
method: uni-directional, sequence direction: sequenced from T3 primer
(5' -> 3')

Project: "Silkworm Genome Program in MAF, and Research for the
Future Program in JSPS". see "SilkBase"
Future Program in JSPS. see "SilkBase"
<http://www.ab.a.u-tokyo.ac.jp/silkbase/>, for whole ESTdb.
Location/Qualifiers
1..531
/organism="Bombyx mori"
/strain="Shuko x Ryuhaku"
/db_xref="taxon:7091"
/clone="pg--0086"
/clone_lib="Bombyx mori pheromone gland Shuko x Ryuhaku
newly-released adult"

FEATURES

Source

```

/sex="female"
/tissue_type="pheromone gland"
/dev_stage="newly-eclosed adult"
BASE COUNT      119 a      134 c      148 g      130 t
ORIGIN

```

Query Match	73.3%	Score 17.6;	DB 10,	length 531,
Best Local Similarity	83.3%	Pred. No. 1e+02;		
Matches 20;	Conservative 0;	Mismatches 4;	Indels 0;	Gaps 0;

```
QY      1 tcgtgaacgttcgttaacgttcg 24
          |||||  |||||  ||
Db      485 TCGTGAACCGCCGTTAACGTC CG 462
```

RESULT	2
LOCUS	AL118705
DEFINITION	AL118705 571 bp mRNA.
ACCESSION	DKE282761J0510.r1.761 (synonym: hamy?) Homo sapiens cDNA clone.
VERSION	AL118705
KEYWORDS	DKE282761J0510.5, mRNA sequence.
SOURCE	AL118705.1 GI:5924604
ORGANISM	EST.
REFERENCE	human.
COMMENTARY	

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.	Blum, H., Baerendsch, S., Mewes, H.W., Gassenhuber, J. and Wiemann, S.	1 (bases 1 to 571)	EST (Blum, et al.)	Unpublished (1999) Contact: Blum, H

Am Klopferspitz 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ), Email s.wiemann@dkfz-heidelberg.de;
sequenced by LMU (Ludwig Maximilians University,
Munich/Germany) within the cDNA sequencing consortium of the German
Genome Project.
No *in situ* sequence available.
This clone (DKRZp761D0510) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de

FEATURES
source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp761L0510"
/clone_lib="761 (synonym: hamy2)"
/tissue_type="amygdala"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"

```

Query Match	73.3%	Score 17.6	DB 10	Length 571
Best Local Similarity	83.3%	Pred. No. 1e+02		
Matches 20	Conservative 0	Mismatches 4	Indels 0	Gaps 0

QY	1	tcgtcgaa	cttcgtta	cgttcg	24
Db	452	TCGACG	GAGCTT	CGTAAC	CGTTG 475

RESULT	3
LOCUS	A0623639
DEFINITION	A0623639 609 bp DNA GSS 16-JUN-1999 HS_5577.A2.F05.SPEE.RPCT.11 Human Male BAC Library Homo sapiens genomic clone Plate=953 Col=10 Row=K, DNA sequence.

ACCESSION	AO623639	GI:5086119
VERSION	AO623639.1	
KEYWORDS	GSS.	
SOURCE	human.	
ORGANISM	Homo sapiens	

REFERENCE

AUTHORS

TITLE Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and Hood, L.
Sequence-tagged connectors: A sequence approach to mapping and

TITLE	Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
JOURNAL	Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
COMMENT	
MDLINE	99380589
CONTACT	Mahairas GG, Wallace JC, Hood L

University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallaceu.washington.edu
Clones are derived from the human BAC library RPC1-11. For BAC library availability, please contact Pieter de Jong (pieterdejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm) or from Resear h Genetics (<http://inforesgen.com>). BAC end Web Server: <http://www.htsc.washington.edu>
Plate: 953 row: K column: 10
Seq primer: SP6
Class: BAC ends
High quality sequence stop: 609.

FEATURES
source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="plate353 Col-10 Row-K"
/clone_id="PC1-11 Human Male BAC Library"
/sex="male"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACe3.6 vector at EcoRI sites"
BASE COUNT      185 a      124 c      162 g      125 t      13 others
ORIGIN

```

Query Match	73.3%	Score 17.6;	DB 13.	Length 609;
Best Local Similarity	83.3%	Pred. No. 1e+02;		
Matches 20;	Conservative 0;	Mismatches 4;	Indels 0;	Gaps 0;

QY	1	tcgtgaacgltcgtaacgttcg	24
Db	335	TCGTCGACGTTGCAATACGTCG	358

RESULT	4
CNS02XJ3	
LOCUS	CNS02XJ3 794 bp DNA
DEFINITION	GSS Tetraodon nigroviridis genome survey sequence PUC-ori end of clone 178F10 of library G from Tetraodon nigroviridis, genomic survey sequence.
ACCESSION	AL218424
VERSION	AL218424.1 GI:7877243
KEYWORDS	GSS; genome survey sequence.
SOURCE	Tetraodon nigroviridis.
ORGANISM	Tetraodon nigroviridis

REFERENCE

1 (bases 1 to 794)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.

AUTHORS Roest-Crollius, H., Jaillon, O., Dasilva, C., Fizames, C., Fisher, C., Bounneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A. and Weissenbach, J.

TITLE Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 794)

AUTHORS Roest-Crollius, H., Jaillon, O., Dasilva, C., Bounneau, L., Fisher, C., Bernot, A., Fizames, C., Wincker, P., Brottier, P., Quetier, F., Saurin, W. and Weissenbach, J.

TITLE Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence

JOURNAL Unpublished

REFERENCE 3 (bases 1 to 794)

AUTHORS Genoscope.

TITLE Direct Submission

JOURNAL Submitted (12-APR-2000) to the EMBL/Genbank/DBJ databases

COMMENT This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/Tetraodon>.

FEATURES Location/Qualifiers

source

1..794

/organism="Tetraodon nigroviridis"

/db_xref="taxon:99883"

/clone_id="178F10"

/note="Genoscope sequence ID : COAG178DC05SP1-end : PUC-ori"

BASE COUNT 162 a 232 c 239 g 161 t

ORIGIN

Query Match 71.7%; Score 17.2; DB 13; Length 794;
Best Local Similarity 86.4%; Pred. No. 1.7e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 gtgcagcttcgttaacgttcg 24
||||| ||||| ||||| ||||| |||||

Db 11 GTCGCTCGTGTAAACGTTCC 32

RESULT 5

LOCUS BF860303 2228 bp mRNA EST 19-JAN-2001

DEFINITION 963016B12.x1 C. reinhardtii CC-1690, Stress condition I, normalized

ACCESSION BF860303

VERSION BF860303.1 GI:12250440

KEYWORDS EST.

SOURCE Chlamydomonas reinhardtii.

ORGANISM Chlamydomonas reinhardtii.
Chlamydomonas reinhardtii.
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadales; Chlamydomonadales.

REFERENCE 1 (bases 1 to 2228)

AUTHORS Grossman, A., Davies, J., Federspiel, N., Harris, E., Hauser, C., Lefebvre, P., Mcdermott, J. P., Shrago, J., Sillflow, C. and Stern, D.

TITLE Analyses of the Chlamydomonas reinhardtii Genome: A Model, Unicellular System for Analyzing Gene Function and Regulation in Vascular Plants; project phase 3

JOURNAL Unpublished (2000)

COMMENT Contact: Charles Hauser
DCMB Box 91000
Duke University
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
Email: chaus@duke.edu.

FEATURES Location/Qualifiers

source

1..2228

/organism="Chlamydomonas reinhardtii"

/strain="CC-1690 wild type mt+ 21gr"

/db_xref="taxon:3055"

/clone_id="C. reinhardtii CC-1690, Stress condition I, normalized, Lambda Zap II"

BASE COUNT 633 a 427 c 543 g 470 t 155 others

ORIGIN

Query Match 71.7%; Score 17.2; DB 11; Length 2228;
Best Local Similarity 79.2%; Pred. No. 2.2e+02;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 tctgcagcttcgttaacgttcg 24
||||| ||||| ||||| ||||| |||||

Db 1744 TCGTCGATGTTCGATGNGTCC 1767

RESULT 6

LOCUS D37092 360 bp mRNA EST 24-OCT-1994

DEFINITION CELK040E3F Yuiji Kohara unpublished cDNA Caenorhabditis elegans cDNA

ACCESSION D37092

VERSION D37092.1 GI:5525474

KEYWORDS EST.

SOURCE Caenorhabditis elegans.

ORGANISM Caenorhabditis elegans.
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae
; Rhabditidae; Pseudocercariae; Caenorhabditis.

REFERENCE 1 (bases 1 to 360)

AUTHORS Kohara, Y., Mitsuk, H., Nishigaki, A., Motohashi, T., Sugimoto, A. and Tabara, H.

TITLE Toward an expression map of the C. elegans genome

JOURNAL Unpublished (1994)

COMMENT Contact: Yuiji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.

FEATURES Location/Qualifiers

source

1..360

/organism="Caenorhabditis elegans"

/strain="CB1489 him-8(c1489)"

/db_xref="taxon:6239"

/clone_id="YK40E3"

/clone_id="Yuiji Kohara unpublished cDNA"

/sex="hermaphrodite, male"

/tissue-type="whole animal"

/dev_stage="varied"

BASE COUNT 95 a 76 c 73 g 115 t 1 others

ORIGIN

Query Match 69.2%; Score 16.6; DB 11; Length 360;
Best Local Similarity 82.6%; Pred. No. 2.9e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 cgtgcagcttcgttaacgttcg 24
||||| ||||| ||||| ||||| |||||

Db 204 CGTCGAATAATCTTAACGATCG 226

RESULT 7
 C12646 379 bp mRNA EST 28-DEC-1998
 LOCUS C12646 yuji.kohara unpublished cDNA Caenorhabditis elegans CDNA
 DEFINITION clone yk153h4 5', mRNA sequence.
 ACCESSION C12646
 VERSION C12646.1 GI:1560199
 KEYWORDS EST.
 SOURCE Caenorhabditis elegans.
 ORGANISM Caenorhabditis elegans.
 Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
 ; Rhabditidae; Pelodicerinae; Caenorhabditis.
 1 (bases 1 to 379)
 Kohara,Y., Mochizuki,T., Tabara,H., Watanabe,H., Sugimoto,A., Sano
 M., Miyata,A. and Nishigaki,A.
 Expression map of the C.elegans genome
 Unpublished (1996)
 JOURNAL Contact: Yuji Kohara
 COMMENT Genome Biology Lab.
 National Institute of Genetics
 Yata 1111, Mishima, Shizuoka 411, Japan
 Tel: 81-559-81-6854
 Fax: 81-559-81-6855
 Email: yk@haraelab.nig.ac.jp.
 Location/Qualifiers
 1..379
 /organism="Caenorhabditis elegans"
 /strain="CB1489 him-8(e1489)"
 /db_xref="taxon:6239"
 /clone="yk153h4"
 /clone_1ib="yuji Kohara unpublished cDNA"
 /sex="hermaphrodite, male"
 /tissue_type="whole animal"
 /dev_stage="varied"
 BASE COUNT 95 a 90 c 94 g 92 t 8 others
 ORIGIN

Query Match 69.2%; Score 16.6; DB 11; Length 379;
 Best Local Similarity 79.2%; Pred. No. 3e+02;
 Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 tcgtgaacgttcgttaacgttcg 24
 ||||| ||||| ||||| ||||| |||||
 Db 115 TCGTGACATTCGNCACATTCG 138

RESULT 8
 A1531538 460 bp mRNA EST 19-APR-2001
 LOCUS SDD2623.3prime SD Drosophila melanogaster Schneider L2 cell culture
 DEFINITION port2 Drosophila melanogaster CDNA clone SDD2623 3prime similar to
 U21123: Drosophila melanogaster ena polypeptide gene, complete cds,
 mRNA sequence.
 ACCESSION A1531538
 VERSION A1531538.1 GI:4445673
 KEYWORDS EST.
 SOURCE fruit fly.
 ORGANISM Drosophila melanogaster.
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 460)
 Harvey,D., Broksstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
 Lewis,S. and Rubin,G.M.
 BDGP/HMT Drosophila EST Project
 Unpublished (2001)
 JOURNAL Contact: Stapleton, M.
 COMMENT BDGP
 Lawrence Berkeley National Lab
 One Cyclotron Rd, Berkeley, CA 94720, USA

Fax: 510 486 6798
 Email: http://www.fruitfly.org/EST_estefruitfly.berkeley.edu
 Based upon the presence of a XhoI site followed by a run of 14 or
 more T residues at the beginning of the sequence, this clone
 probably contains an inverted insert. The resulting Poly-T sequence
 has been removed.
 Plate: 26 row: B column: 11
 High quality sequence stop: 268.
 Location/Qualifiers
 1..460
 /organism="Drosophila melanogaster"
 /db_xref="taxon:7227"
 /clone="SDD2623"
 /clone_1ib="SD Drosophila melanogaster Schneider L2 cell
 culture port2"
 /lab_host="DH5-alpha"
 /note="Vector: port2; Site_1: EcoRI; Site_2: XhoI; Sized
 fractionated cDNAs were directly ligated into port2.
 Plasmid cDNA library."
 BASE COUNT 106 a 119 c 96 g 139 t
 ORIGIN

Query Match 69.2%; Score 16.6; DB 10; Length 460;
 Best Local Similarity 82.6%; Pred. No. 3.1e+02;
 Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 tcgtgaacgttcgttaacgttcg 23
 ||||| ||||| ||||| ||||| |||||
 Db 80 TCGTGATCGATCGACATTCG 102

RESULT 9
 A0860867/c 464 bp DNA GSS 03-NOV-1999
 LOCUS nbed0015104r CUI1 Rice BAC Library (EcoRI) Oryza sativa genomic
 DEFINITION clone nbed0015104r, DNA sequence.
 ACCESSION A0860867
 VERSION A0860867.1 GI:6211324
 KEYWORDS GSS.
 SOURCE Oryza sativa.
 ORGANISM Oryza sativa.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.
 1 (bases 1 to 464)
 Wing,R.A. and Dean,R.A.
 A BAC End Sequencing Framework to Sequence the Rice Genome
 Unpublished (1998)
 JOURNAL Contact: Wing RA
 COMMENT Clemson University Genomics Institute
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 7288
 Fax: 864 656 4293
 Email: rwing@clemson.edu
 Seg primer: GGAACACGCTATGACCATG
 Class: BAC ends
 High quality sequence start: 22
 High quality sequence stop: 363.
 Location/Qualifiers
 1..464
 /organism="Oryza sativa"
 /strain="Japonica"
 /cultivar="Nipponbare"
 /db_xref="taxon:4530"
 /clone="nbed0015104r"
 /clone_1ib="CUI1 Rice BAC Library (EcoRI)"
 /tissue_type="Leaf"
 /lab_host="E. coli DH10B"
 /note="Vector: pBACindigo; site_1: EcoRI; site_2: EcoRI;
 Rice is the most important food crop in the world. Half of
 the world population, especially those inhabiting highly

```

/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="RD3655"
/clone_lib="LD Drosophila melanogaster embryo POT2"
/sex="male and female"
/dev_stage="0 to 24 hours mixed stage embryonic"
/lab_host="Xyl Blue"

```

```

/organism="Mus musculus"
/strain="C3H"
/db_xref="taxon:10090"
/clone="IMAGE:902305"
/clone_1lp="Barstread mouse myotubes MFLRB5"
/cell_line="C2C12"
/lab_host="DH10B"
/notes="vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA
was primed with a Not I - oligo(dt) primer [5',
TGTATACGATCTCTGAAGTGGAGCGCGCCCTCTTTTTTTTTTTTTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
[AAATCGATCCCTGG], digested with Not I and cloned into the
Not I and Eco RI sites of the modified pT7T3 vector.
Library constructed by Bob Barstread. The C2C12 cell line
(available from ATCC, catalog # CRL-1772) differentiates
rapidly, forming contractile myocubes and producing
characteristic muscle proteins."

```

Query Match 69.2%; Score 16.6; DB 10; Length 522;
 Best Local Similarity 82.6%; Pred. No. 3.2e+02;
 Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 2 cgtcgaacgttcgttaacgttc 24
 ||||| ||||| ||||| ||||| |||||
 Db 148 CGCCGAAGTGTGACGCTCC 126

RESULT 12
 Bg638934 532 bp mRNA EST 23-APR-2001
 LOCUS LD38917.3prime LD Drosophila melanogaster embryo POT2 Drosophila
 DEFINITION melanogaster cDNA clone LD38917.3 similar to enh: FBan0015112
 'signal transduction' located on: 2R 56B5-56B5:: 04/10/2001, mRNA
 sequence.
 ACCESSION Bg638934
 VERSION Bg638934
 KEYWORDS EST.
 SOURCE fruit fly
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 REFERENCE 1 (bases 1 to 532)
 AUTHORS Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G.,
 Lewis, S., and Rubin, G.M.
 BDCP/HMI Drosophila EST Project
 TITLE Unpublished (2001)
 JOURNAL Other ESTs: LD38917.5prime
 COMMENT Contact: Stapleton, M.
 BDCP
 Lawrence Berkeley National Lab
 One Cyclotron Rd, Berkeley, CA 94720, USA
 Fax: 510 486 6798
 Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
 Based upon the presence of a XhoI site followed by a run of 14 or
 more T residues at the beginning of the sequence, this clone was
 polyadenylated. The resulting Poly-T sequence has been removed. hit
 genome AE003797: arm:2R [13841036,14128894]
 estimated-cyto:55F6-56C1: 04/10/2001
 Plate: LD.389 row: B column: 5
 High quality sequence stop: 504.

FEATURES
 Location/Qualifiers
 1..532
 /organism="Drosophila melanogaster"
 /db_xref="taxon:7227"
 /clone="LD38917"
 /clone_lib="LD Drosophila melanogaster embryo POT2"
 /sex="male and female"
 /dev_stage="0 to 24 hours mixed stage embryonic"
 /lab_host="XLI Blue"
 /note="Organ: embryo; Vector: POT2; Site_1: EcoRI; Site_2:
 XhoI; Sized fractionated cDNAs were directly ligated into
 POT2."

BASE COUNT 113 a 147 c 112 g 160 t
 ORIGIN

Query Match 69.2%; Score 16.6; DB 11; Length 532;
 Best Local Similarity 82.6%; Pred. No. 3.2e+02;
 Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 tcgtcgaacgttcgttaacgttc 23
 ||||| ||||| ||||| ||||| |||||
 Db 80 TCGTCGATCGATCGTCAACATTC 102

RESULT 13
 A1976120 537 bp mRNA EST 27-AUG-1999
 LOCUS A1976120

DEFINITION EST720714 Schistosoma mansoni female, Phil Loverde/Joe Merrick
 Schistosoma mansoni cDNA clone SMFB14.5' end, mRNA sequence.
 ACCESSION A1976120
 VERSION A1976120.1 GI:5789288
 KEYWORDS EST.
 SOURCE Schistosoma mansoni.
 ORGANISM Schistosoma mansoni
 Eukaryota; Metazoa; Platyhelminthes; Turbellarian Platyhelminths;
 Rhabditophora; Euliothophora; Revertepermatia; Mediofusata;
 Neodermata; Trematoda; Digenea; Strigoidia; Schistosomatidae;
 Schistosomatidae; Schistosoma.
 REFERENCE 1 (bases 1 to 537)
 AUTHORS Merrick, J.M., Osman, A., Loverde, P.T., Chandra, I., Glodex, A., Fraser,
 C.M., and Lee, N.H.
 TITLE Generation of a Schistosoma sp. EST (SMEST) Catalog & Schistosoma
 Gene Index
 JOURNAL Unpublished (1998)
 COMMENT Contact: Norman H. Lee
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 3529
 Fax: 301 838 0208
 Email: nhlee@tigr.org
 Seq primer: M13 Reverse.
 Location/Qualifiers
 1..537
 /organism="Schistosoma mansoni"
 /db_xref="taxon:6183"
 /clone="SMFB14"
 /clone_lib="Schistosoma mansoni female, Phil Loverde/Joe
 Merrick"
 /sex="female"
 /note="Vector: pBluescript SK; Site_1: EcoRI; Site_2: XhoI
 ; directionally cloned cDNAs"

BASE COUNT 162 a 99 c 125 g 151 t
 ORIGIN

Query Match 69.2%; Score 16.6; DB 10; Length 537;
 Best Local Similarity 82.6%; Pred. No. 3.2e+02;
 Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 tcgtcgaacgttcgttaacgttc 23
 ||||| ||||| ||||| ||||| |||||
 Db 268 TCGTCGAACGTTGATTAACGCTCC 290

RESULT 14
 Bg795168 552 bp mRNA EST 16-MAY-2001
 LOCUS UTSW_SM266 UTSW Adult Mouse Skeletal Muscle Library Mus musculus
 DEFINITION UTSW_SM266 UTSW Adult Mouse Skeletal Muscle Library Mus musculus
 Bg795168
 Bg795168.1 GI:14130738
 ACCESSION Bg795168
 VERSION Bg795168.1 GI:14130738
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 552)
 AUTHORS Gallardo, T.D., Schageman, J.J., Pertsemilds, A., Garner, H.R.,
 Williams, R.S., and Shohet, R.V.
 TITLE UTSW Southwestern Medical Center, Adult Mouse Skeletal Muscle cDNA
 Library
 JOURNAL Unpublished (2001)
 COMMENT Contact: Schageman, J.J.
 Shohet/Garner labs
 University of Texas Southwestern Medical Center
 6000 Harry Hines Blvd., Nk2.226, Dallas, TX 75390, USA
 Tel: 214 648 1674
 Email: Jeff.Schageman@UTSouthwestern.edu
 cDNA library constructed by UTSW as a component of the program for
 Genomic Applications (PGA) and the Reynolds Heart Disease

Prevention grants for use in cDNA microarray experiments. Sequence Quality: Sequence ends were trimmed based on percentage of ambiguous base calls or 'N's in windowed segments. Sequencing: First-pass sequencing; ABI Prism 377 sequencer and analysis software. Seq primer: M13/pUC Reverse.

FEATURES

Location/Qualifiers

```
1..552
  /organism="Mus musculus"
  /db_xref="taxon:10090"
  /clone="UTSM_SM2E6"
  /clone_lib="UTSM Adult Mouse Skeletal Muscle Library"
  /sex="Pooled"
  /tissue_type="Diaphragm/Hind limb muscles"
  /cell_type="Skeletal muscle"
  /dev_stage="2 months"
  /lab_host="DH5a"
  /note="Vector: pAMP10 (Gibco): Cloned unidirectionally.
  Primer: Oligo dT. RNA Isolation: cytoplasmic RNA preps
  (Maniatis); Cloning Technique: CDA Cloning (CloneAmp,
  Life Technologies); Average Insert size: 1.8 kb;
  Insertion site: TACGCCACTGATTCGAGTC-->. Other
  information regarding entire library may be found at
  http://pga.swmed.edu/Data/Libraries/microarray_cdna_library.htm."
```

BASE COUNT 156 a 133 c 157 g 104 t 2 others

ORIGIN

Query Match 69.2%; Score 16.6; DB 11; Length 552;
Best Local Similarity 82.6%; Pred. No. 3.2e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 2 cgtcgaacgttcgttaacgttcg 24
||||| ||||| ||||| ||||| |||||
Db 483 CGTCCACTTGTGTTAACGTTGG 461

RESULT 15 576 bp mRNA EST 21-MAY-2001
BG799488/c fms97f06.y1 zebrafish Research Genetics C32 fin Danio rerio cDNA
LOCUS clone 4468762 5' similar to SW:VASP_HUMAN P50552
DEFINITION VASODILATOR-STIMULATED PHOSPHOPROTEIN;; mRNA sequence.

ACCESSION BG799488
VERSION BG799488.1 GI:14163820
KEYWORDS EST.

SOURCE

ORGANISM zebrafish.
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Rasbora; Danio.

REFERENCE

AUTHORS Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Maria,M., Eddy,
,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood
,R., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,
Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E.,
Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.
and Wilson,R.

TITLE

Mashu zebrafish EST Project 1998
Unpublished (1998)
Other-ESTs: fms97f06.x1

JOURNAL

COMMENT

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: zbrafish@wustl.edu
CDNA Library Preparation: Ning Wu. CDNA Library Arrayed by: Steve
Johnson. DNA Sequencing by: Washington University Genome Sequencing
Center. Clone distribution: Research Genetics web address:
http://www.researchgenetics.com/
zebrafish identity (P-value greater than 1e-99) found to:
g112289350|db|AA542615|AA542615 fa08b01.r1 zebrafish ICRFzfls Danio

rerio cDNA
High quality sequence stop: 489.
Location/Qualifiers

FEATURES

1..576
/organism="Danio rerio"
/db_xref="taxon:7955"

```
/clone="4468762"
/clone_lib="zebrafish Research Genetics C32 fin"
/tissue_type="Fin"
/lab_host="GeneHogs (HS996, a phage-resistant isolate of
DH10B)"
/note="Vector: pT73D-Pac with a modified polylinker;
Site_1: EcoRI; Site_2: NotI; 1st strand cDNA was prepared
from zebrafish(C32) fin, and was then primed with a Not I
- oligo(dT) primer. Double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pT73 vector. Library is non-normalized. Library was
constructed by Ning Wu. NOTE: Clones from this library are
only available thru Research Genetics (www.resgen.com)."
```

BASE COUNT 146 a 166 c 146 g 118 t

ORIGIN

Query Match 69.2%; Score 16.6; DB 11; Length 576;
Best Local Similarity 82.6%; Pred. No. 3.2e+02;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 tcgtcgaacgttcgttaacgttc 23
||||| ||||| ||||| ||||| |||||
Db 28 TCGCCGTCGTCGCTTAACGTTTC 6

Search completed: March 11, 2002, 20:30:08
Job time: 2946 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2002, 21:01:45 ; Search time 100.2 Seconds
(without alignments)
54.246 Million cell updates/sec

Title: US-09-802-518-9

Perfect score: 24

Sequence: 1 tcgtcgaacgttcgttaacgttcg 24

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA:*

1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*

2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*

3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*

4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*

5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq:*

6: /cgn2_6/ptodata/1/ina/backfillseq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
C 1	16.4	68.3	2959 1 US-08-408-188A-1	Sequence 1, Appli
C 2	16.4	68.3	2959 1 US-08-408-188A-2	Sequence 2, Appli
C 3	16.4	68.3	2959 1 US-08-408-188A-3	Sequence 3, Appli
C 4	16.4	68.3	2959 1 US-08-408-188A-4	Sequence 4, Appli
C 5	16.4	68.3	2959 1 US-08-766-488-1	Sequence 1, Appli
C 6	16.4	68.3	2959 1 US-08-766-488-2	Sequence 2, Appli
C 7	16.4	68.3	2959 1 US-08-766-488-3	Sequence 3, Appli
C 8	16.4	68.3	2959 1 US-08-766-488-4	Sequence 4, Appli
C 9	14.8	61.7	447 1 US-08-555-386-4	Sequence 4, Appli
10	14.8	61.7	4529 1 US-08-555-386-1	Sequence 1, Appli
11	14.6	60.8	1303 1 US-08-844-154-1	Sequence 1, Appli
12	14.6	60.8	1303 2 US-09-126-192A-1	Sequence 1, Appli
13	14.6	60.8	1344 2 US-08-902-585-1	Sequence 1, Appli
C 14	14.6	60.8	1777 3 US-08-961-083-221	Sequence 221, App
C 15	14.6	60.8	1879 3 US-08-961-083-219	Sequence 219, App
16	14.6	60.8	2096 5 PCT-US94-08120-9	Sequence 9, Appli
17	14.6	60.8	2096 5 PCT-US94-08120-9	Sequence 9, Appli
18	14.6	60.8	2096 5 PCT-US94-12913A-9	Sequence 9, Appli
19	14.6	60.8	2099 1 US-08-094-533B-9	Sequence 9, Appli
20	14.6	60.8	2099 1 US-08-276-860A-9	Sequence 9, Appli
21	14.6	60.8	2099 1 US-08-444-393-9	Sequence 9, Appli
22	14.6	60.8	2099 1 US-08-799-913-9	Sequence 9, Appli
23	14.6	60.8	2099 2 US-08-711-893-9	Sequence 9, Appli
24	14.6	60.8	2099 2 US-09-150-200-9	Sequence 9, Appli
25	14.6	60.8	2099 3 US-09-150-201-9	Sequence 9, Appli
26	14.6	60.8	2099 4 US-09-452-370-9	Sequence 9, Appli
27	14.6	60.8	3622 1 US-08-306-691B-13	Sequence 13, Appli

ALIGNMENTS

28	14.6	60.8	3622 5 PCT-US93-06251-24	Sequence 24, Appli
C 29	14.4	60.0	1294 1 US-07-662-005A-3	Sequence 3, Appli
30	14.4	60.0	1611 1 US-08-485-569-1	Sequence 1, Appli
31	14.4	60.0	1611 1 US-08-480-993-1	Sequence 1, Appli
32	14.4	60.0	1611 2 US-07-903-079B-1	Sequence 1, Appli
33	14.4	60.0	1836 1 US-07-754-918A-1	Sequence 1, Appli
C 34	14.4	60.0	1850 3 US-08-617-860B-32	Sequence 32, Appli
C 35	14.4	60.0	1959 3 US-08-987-123-1	Sequence 1, Appli
C 36	14.4	60.0	1959 3 US-08-987-123-3	Sequence 3, Appli
37	14.4	60.0	2894 1 US-08-278-091-1	Sequence 1, Appli
38	14.4	60.0	2894 1 US-08-483-859-1	Sequence 1, Appli
39	14.4	60.0	2894 1 US-08-472-173-1	Sequence 1, Appli
40	14.4	60.0	2894 2 US-08-487-167-1	Sequence 1, Appli
41	14.4	60.0	2894 2 US-08-482-816-1	Sequence 1, Appli
42	14.4	60.0	2894 2 US-08-296-149-1	Sequence 1, Appli
43	14.4	60.0	2894 2 US-08-801-499-1	Sequence 1, Appli
44	14.4	60.0	2894 2 US-08-615-271-1	Sequence 1, Appli
45	14.4	60.0	2894 3 US-09-074-660-1	Sequence 1, Appli

RESULT 1
US-08-408-188A-1/c
Sequence 1, Application US/08408188A
Patent No. 5616480
GENERAL INFORMATION:
APPLICANT: SUGIMOTO, Masakazu
APPLICANT: KOJIMA, Hiroyuki
APPLICANT: TANAKA, Akiko
APPLICANT: MATSUI, Hiroshi
APPLICANT: SATOH, Katsunaki
APPLICANT: NAKAMATSU, Tsuyoshi
TITLE OF INVENTION: TEMPERATURE-SENSITIVE PLASMID
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: ORION, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/408,188A
FILING DATE: 22-MAR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/182,361
FILING DATE: 18-JAN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/774,846
FILING DATE: 11-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-273348
FILING DATE: 15-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 3-245291
FILING DATE: 19-JUN-1991
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 10-522-0 FWC I
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 2959 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Corynebacterium glutamicum
STRAIN: ATCC 13058
US-08-408-188A-1

Query Match 68.3%; Score 16.4; DB 1; Length 2959;
Best Local Similarity 94.4%; Pred. No. 12;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 cgaacgttcgtaacgtt 22
|||||
DB 432 CGACCTTCGTTAAAGTT 415

RESULT 2

US-08-408-188A-2/c
Sequence 2, Application US/08408188A
Patent No. 5616480
GENERAL INFORMATION:
APPLICANT: SUGIMOTO, Masakazu
APPLICANT: KOJIMA, Hiroyuki
APPLICANT: TANAKA, Akiko
APPLICANT: MATSUI, Hiroshi
APPLICANT: SATOH, Katsunaki
APPLICANT: NAKAMATSU, Tsuyoshi
TITLE OF INVENTION: TEMPERATURE-SENSITIVE PLASMID
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT,
P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/408,188A
FILING DATE: 22-MAR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/182,361
FILING DATE: 18-JAN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/774,846
FILING DATE: 11-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-273348
FILING DATE: 15-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 3-245291
FILING DATE: 19-JUN-1991
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 10-522-0 FWC I
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2959 base pairs

TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Corynebacterium glutamicum
STRAIN: ATCC 13058
US-08-408-188A-2

Query Match 68.3%; Score 16.4; DB 1; Length 2959;
Best Local Similarity 94.4%; Pred. No. 12;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 cgaacgttcgtaacgtt 22
|||||
DB 432 CGACCTTCGTTAAAGTT 415

RESULT 3

US-08-408-188A-3/c
Sequence 3, Application US/08408188A
Patent No. 5616480
GENERAL INFORMATION:
APPLICANT: SUGIMOTO, Masakazu
APPLICANT: KOJIMA, Hiroyuki
APPLICANT: TANAKA, Akiko
APPLICANT: MATSUI, Hiroshi
APPLICANT: SATOH, Katsunaki
APPLICANT: NAKAMATSU, Tsuyoshi
TITLE OF INVENTION: TEMPERATURE-SENSITIVE PLASMID
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT,
P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/408,188A
FILING DATE: 22-MAR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/182,361
FILING DATE: 18-JAN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/774,846
FILING DATE: 11-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-273348
FILING DATE: 15-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 3-245291
FILING DATE: 19-JUN-1991
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 10-522-0 FWC I
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2959 base pairs
STRANDEDNESS: double

TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Corynebacterium glutamicum
STRAIN: ATCC 13058
US-08-408-188A-3

Query Match 68.3%; Score 16.4; DB 1; Length 2959;
Best Local Similarity 94.4%; Pred. No. 12;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Y 5 cgaacgttcgttaacgtt 22
|||||
Db 432 CGAACGTTCTTAAGTT 415

RESULT 4
US-08-408-188A-4/C
Sequence 4, Application US/08408188A
Patent No. 5616480
GENERAL INFORMATION:
APPLICANT: SUGIMOTO, Masakazu
APPLICANT: KOJIMA, Hiroyuki
APPLICANT: TANAKA, Akiko
APPLICANT: MATSUI, Hiroshi
APPLICANT: SATOH, Katsuki
APPLICANT: NAKAMATSU, Tsuyoshi
TITLE OF INVENTION: TEMPERATURE-SENSITIVE PLASMID
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLOH, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESS: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/408,188A
CLASSIFICATION: 435
FILING DATE: 22-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/182,361
FILING DATE: 18-JAN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/774,846
FILING DATE: 11-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-273348
FILING DATE: 15-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 3-245291
FILING DATE: 19-JUN-1991
ATTORNEY/AGENT INFORMATION:
NAME: OBLOH, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 10-522-0 FWC I
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2959 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

ORIGINAL SOURCE:
ORGANISM: Corynebacterium glutamicum
STRAIN: ATCC 13058
US-08-408-188A-4

Query Match 68.3%; Score 16.4; DB 1; Length 2959;
Best Local Similarity 94.4%; Pred. No. 12;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Y 5 cgaacgttcgttaacgtt 22
|||||
Db 432 CGAACGTTCTTAAGTT 415

RESULT 5
US-08-766-488-1/C
Sequence 1, Application US/08766488
Patent No. 5756347
GENERAL INFORMATION:
APPLICANT: SUGIMOTO, Masakazu
APPLICANT: KOJIMA, Hiroyuki
APPLICANT: TANAKA, Akiko
APPLICANT: MATSUI, Hiroshi
APPLICANT: SATOH, Katsuki
APPLICANT: NAKAMATSU, Tsuyoshi
TITLE OF INVENTION: TEMPERATURE-SENSITIVE PLASMID
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLOH, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESS: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/766,488
CLASSIFICATION: 435
FILING DATE: 13-DEC-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/408,188
FILING DATE: 22-MAR-1995
APPLICATION NUMBER: US 08/182,361
FILING DATE: 18-JAN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/774,846
FILING DATE: 11-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-273348
FILING DATE: 15-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 3-245291
FILING DATE: 19-JUN-1991
ATTORNEY/AGENT INFORMATION:
NAME: OBLOH, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 10-522-0 FWC I
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2959 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

ORIGINAL SOURCE:
ORGANISM: Corynebacterium glutamicum
STRAIN: ATCC 13058
US-08-766-488-1

Query Match 68.3%; Score 16.4; DB 1; Length 2959;
Best Local Similarity 94.4%; Pred. No. 12;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 cgaacgttcgttaacgtt 22
|||||
DB 432 CGACGCTGCTTAAGTT 415

RESULT 6
US-08-766-488-2/c
Sequence 2, Application US/08766488
Patent No. 5756347
GENERAL INFORMATION:
APPLICANT: SUGIMOTO, Masakazu
APPLICANT: KOJIMA, Hiroyuki
APPLICANT: TANAKA, Akiko
APPLICANT: MATSUI, Hiroshi
APPLICANT: SATOH, Katsunaki
APPLICANT: NAKAMATSU, Tsuyoshi
TITLE OF INVENTION: TEMPERATURE-SENSITIVE PLASMID
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT,
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/766,488
FILING DATE: 13-DEC-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/408,188
FILING DATE: 22-MAR-1995
APPLICATION NUMBER: US 08/182,361
FILING DATE: 18-JAN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/774,846
FILING DATE: 11-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-273348
FILING DATE: 15-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 3-245291
FILING DATE: 19-JUN-1991
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 10-522-0 FWC I
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2959 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

ORIGINAL SOURCE:
ORGANISM: Corynebacterium glutamicum
STRAIN: ATCC 13058
US-08-766-488-2

Query Match 68.3%; Score 16.4; DB 1; Length 2959;
Best Local Similarity 94.4%; Pred. No. 12;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 cgaacgttcgttaacgtt 22
|||||
DB 432 CGACGCTGCTTAAGTT 415

RESULT 7
US-08-766-488-3/c
Sequence 3, Application US/08766488
Patent No. 5756347
GENERAL INFORMATION:
APPLICANT: SUGIMOTO, Masakazu
APPLICANT: KOJIMA, Hiroyuki
APPLICANT: TANAKA, Akiko
APPLICANT: MATSUI, Hiroshi
APPLICANT: SATOH, Katsunaki
APPLICANT: NAKAMATSU, Tsuyoshi
TITLE OF INVENTION: TEMPERATURE-SENSITIVE PLASMID
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT,
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/766,488
FILING DATE: 13-DEC-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/408,188
FILING DATE: 22-MAR-1995
APPLICATION NUMBER: US 08/182,361
FILING DATE: 18-JAN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/774,846
FILING DATE: 11-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-273348
FILING DATE: 15-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 3-245291
FILING DATE: 19-JUN-1991
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 10-522-0 FWC I
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2959 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

ORIGINAL SOURCE:
ORGANISM: Corynebacterium glutamicum
STRAIN: ATCC 13058
US-08-766-488-3

Query Match 68.3%; Score 16.4; DB 1; Length 2959;
Best Local Similarity 94.4%; Pred. No. 12;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 cgaacgttcgtaacgtt 22
|||||
DB 432 CGAAGCTTCGTTAAAGTT 415

RESULT 8
US-08-766-488-4/C
Sequence 4, Application US/08766488
Patent No. 5756347
GENERAL INFORMATION:
APPLICANT: SUGIMOTO, Masakazu
APPLICANT: KOJIMA, Hiroyuki
APPLICANT: TANAKA, Akiko
APPLICANT: MATSUI, Hiroshi
APPLICANT: SATOH, Katsunaki
APPLICANT: NAKAMATSU, Tsuyoshi
TITLE OF INVENTION: TEMPERATURE-SENSITIVE PLASMID
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
ADDRESS: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/766,488
FILING DATE: 13-DEC-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/408,188
FILING DATE: 22-MAR-1995
APPLICATION NUMBER: US 08/182,361
FILING DATE: 18-JAN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/774,846
FILING DATE: 11-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-273348
FILING DATE: 15-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 3-245291
FILING DATE: 19-JUN-1991
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 10-522-0 FMC I
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2959 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

ORIGINAL SOURCE:
ORGANISM: Corynebacterium glutamicum
STRAIN: ATCC 13058
US-08-766-488-4

Query Match 68.3%; Score 16.4; DB 1; Length 2959;
Best Local Similarity 94.4%; Pred. No. 12;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 cgaacgttcgtaacgtt 22
|||||
DB 432 CGAAGCTTCGTTAAAGTT 415

RESULT 9
US-08-565-386-4
Sequence 4, Application US/08565386
Patent No. 5741697
GENERAL INFORMATION:
APPLICANT: Bavoli, Patrick M.
APPLICANT: Hsia, Ru-ching
TITLE OF INVENTION: BACTERIOPHAGE OF CHLAMYDIA PSITTACI
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
STREET: Clinton Square, P.O. Box 1051
CITY: Roches
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/565,386
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Timian, Susan J.
REGISTRATION NUMBER: 34,103
REFERENCE/DOCKET NUMBER: 176/60040
TELECOMMUNICATION INFORMATION:
TELEPHONE: 716-263-1636
TELEFAX: 716-263-1600
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 447 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
US-08-565-386-4

Query Match 61.7%; Score 14.8; DB 1; Length 447;
Best Local Similarity 88.9%; Pred. No. 65;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 cgaacgttcgtaacgtt 22
|||||
DB 248 CGAAGCTTCGTGAAGTT 265

RESULT 10
US-08-565-386-1
Sequence 1, Application US/08565386
Patent No. 5741697
GENERAL INFORMATION:
APPLICANT: Bavoli, Patrick M.
APPLICANT: Hsia, Ru-ching
TITLE OF INVENTION: BACTERIOPHAGE OF CHLAMYDIA PSITTACI
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:

ADDRESSEE: Nixon, Hairgrave, Devans & Doyle LLP
STREET: Clinton Square, P.O. Box 1051
CITY: Roches
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/565,386
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Timian, Susan J.
REGISTRATION NUMBER: 34,103
REFERENCE/DOCKET NUMBER: 176/60040
TELECOMMUNICATION INFORMATION:
TELEPHONE: 716-263-1636
TELEFAX: 716-263-1600
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4529 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
US-08-565-386-1

Query Match 61.7%; Score 14.8; DB 1; Length 4529;
Best Local Similarity 88.9%; Pred. No. 78;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 5 cgaacgtcgttaacgtt 22
||||| ||||| |||||
Db 2731 CCAAACTGCTGACCTT 2748

RESULT 11
US-08-844-154-1
Sequence 1, Application US/08844154
Patent No. 5827708
GENERAL INFORMATION:
APPLICANT: Lawlor, Elizabeth
TITLE OF INVENTION: No. 5827708e1 Compounds
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Smithkline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/844,154
FILING DATE: 18-APR-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 9607993.4
FILING DATE: 18-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimm, Edward R.
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P31457-3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1303 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
US-08-844-154-1

Query Match 60.8%; Score 14.6; DB 1; Length 1303;
Best Local Similarity 81.0%; Pred. No. 89;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 4 tcgaacgtcgttaacgttcg 24
||||| ||||| |||||
Db 1256 TCGAACGTATGTAACCTTCG 1276

RESULT 12
US-09-126-192A-1
Sequence 1, Application US/09126192A
Patent No. 6040163
GENERAL INFORMATION:
APPLICANT: Lawlor, Elizabeth
TITLE OF INVENTION: No. 6040163e1 Compounds
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Smithkline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/126,192A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/844,154
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Gimm, Edward R.
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P31457-3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1303 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
US-09-126-192A-1

Query Match 60.8%; Score 14.6; DB 3; Length 1303;
Best Local Similarity 81.0%; Pred. No. 89;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 4 tcgaacgtcgttaacgttcg 24
||||| ||||| |||||
Db 1256 TCGAACGTATGTAACCTTCG 1276

RESULT 13
US-08-902-585-1
Sequence 1, Application US/08902585
Patent No. 5834276
GENERAL INFORMATION:
APPLICANT: Lawlor, Elizabeth
TITLE OF INVENTION: No. 5834276el Asparaginyl tRNA Synthetase
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/902,585
FILING DATE: 29-JUL-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/844,154
FILING DATE: 18-APR-1997
APPLICATION NUMBER: 9607993.4
FILING DATE: 18-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimmi, Edward R.
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P31457-3/1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1344 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
US-08-902-585-1

Query Match 60.8%; Score 14.6; DB 2; Length 1344;
Best Local Similarity 81.0%; Pred. No. 89;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 4 tcgaacgtcgttaacgttcg 24
||||||| | ||||| |||||
Db 1256 TCGAACGTATGTGTAACCTTCG 1276

RESULT 14
US-08-961-083-221/c
Sequence 221, Application US/08961083
Patent No. 6159469
GENERAL INFORMATION:
APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage

COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,083
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 221:
SEQUENCE CHARACTERISTICS:
LENGTH: 1777 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-961-083-221

Query Match 60.8%; Score 14.6; DB 3; Length 1777;
Best Local Similarity 81.0%; Pred. No. 91;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 2 cgtcgaacgtcgttaacgtt 22
||||||| | ||||| |||||
Db 65 CGTCGAATGTACGTGAACCTT 45

RESULT 15
US-08-961-083-219/c
Sequence 219, Application US/08961083
Patent No. 6159469
GENERAL INFORMATION:
APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,083
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 219:
SEQUENCE CHARACTERISTICS:
LENGTH: 1879 base pairs
TYPE: nucleic acid

; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-961-083-219

Query Match 60.8%; Score 14.6; DB 3; Length 1879;
Best Local Similarity 81.0%; Pred. No. 92;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 cgtcgacgttcgttaacgtt 22
||||||| ||||| |||||
Db 1865 CGTCGATGTACGTGACCTT 1845

Search completed: March 11, 2002, 21:01:48
Job time: 4236 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2002, 20:59:44 ; Search time 227.21 Seconds

(without alignments)
90.559 Million cell updates/sec

Title: US-09-802-518-9

Perfect score: 1 tctgcgaacttcgttaactgctg 24

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapect 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N.Genesec_1101:*

1: /SIDSL/gcgdata/geneseq/geneseqn/NA1980.DAT:*

2: /SIDSL/gcgdata/geneseq/geneseqn/NA1981.DAT:*

3: /SIDSL/gcgdata/geneseq/geneseqn/NA1982.DAT:*

4: /SIDSL/gcgdata/geneseq/geneseqn/NA1983.DAT:*

5: /SIDSL/gcgdata/geneseq/geneseqn/NA1984.DAT:*

6: /SIDSL/gcgdata/geneseq/geneseqn/NA1985.DAT:*

7: /SIDSL/gcgdata/geneseq/geneseqn/NA1986.DAT:*

8: /SIDSL/gcgdata/geneseq/geneseqn/NA1987.DAT:*

9: /SIDSL/gcgdata/geneseq/geneseqn/NA1988.DAT:*

10: /SIDSL/gcgdata/geneseq/geneseqn/NA1989.DAT:*

11: /SIDSL/gcgdata/geneseq/geneseqn/NA1990.DAT:*

12: /SIDSL/gcgdata/geneseq/geneseqn/NA1991.DAT:*

13: /SIDSL/gcgdata/geneseq/geneseqn/NA1992.DAT:*

14: /SIDSL/gcgdata/geneseq/geneseqn/NA1993.DAT:*

15: /SIDSL/gcgdata/geneseq/geneseqn/NA1994.DAT:*

16: /SIDSL/gcgdata/geneseq/geneseqn/NA1995.DAT:*

17: /SIDSL/gcgdata/geneseq/geneseqn/NA1996.DAT:*

18: /SIDSL/gcgdata/geneseq/geneseqn/NA1997.DAT:*

19: /SIDSL/gcgdata/geneseq/geneseqn/NA1998.DAT:*

20: /SIDSL/gcgdata/geneseq/geneseqn/NA1999.DAT:*

21: /SIDSL/gcgdata/geneseq/geneseqn/NA2000.DAT:*

22: /SIDSL/gcgdata/geneseq/geneseqn/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
C 1	16.4	68.3	2958	13	AAO25266
C 2	16.4	68.3	2958	13	AAO25263
C 3	16.4	68.3	2958	13	AAO25265
C 4	16.4	68.7	4142	20	AAK84319
C 5	16.4	66.7	4414	21	AAK47124
C 6	16.4	66.7	13732	21	AAK81484
C 7	16.4	66.7	349980	21	AAK21611
C 8	15.8	65.8	506	20	AAV86097
C 9	15.8	65.8	9619	20	AAK13134
C 10	15.6	65.0	555	22	AAO20845
C 11	15.6	65.0	608	18	AAV75360

12	15.6	65.0	1436	21	AAK64784	Lemon alcohol acyl
C 13	15.6	65.0	3600	21	AAA98970	S. aureus mpf DNA
C 14	15.4	64.2	20	21	AAA96274	Sequence of a stab
C 15	15.4	64.2	20	21	AAA96280	Sequence of a stab
C 16	15.4	62.5	468	21	AAK11680	Aspergillus niger
C 17	15.4	62.5	506	21	AAK44327	Arabidopsis thalia
C 18	15.4	62.5	738	22	AAH65227	C glutamicum codin
C 19	15.4	62.5	1089	22	AAH53863	S. epidermidis ope
C 20	15.4	62.5	1188	21	AAA50733	Yeast type II topo
C 21	15.4	62.5	1320	20	AAK78298	R. capsulatus hema
C 22	15.4	62.5	1479	21	AAK39143	Arabidopsis thalia
C 23	15.4	62.5	1574	21	AAK45383	Arabidopsis thalia
C 24	15.4	62.5	1579	21	AAK33969	Arabidopsis thalia
C 25	15.4	62.5	1584	21	AAK49134	Arabidopsis thalia
C 26	15.4	62.5	1635	20	AAK99510	Nucleic acid sequ
C 27	15.4	62.5	1767	21	AAK32958	Arabidopsis thalia
C 28	15.4	62.5	1769	21	AAK49141	Arabidopsis thalia
C 29	15.4	62.5	1947	22	AAH81338	Escherichia coli p
C 30	15.4	62.5	3603	22	AAH54113	S. epidermidis gen
C 31	15.4	62.5	3757	22	AAH54828	S. epidermidis gen
C 32	15.4	62.5	3816	22	AAH54850	S. epidermidis gen
C 33	15.4	62.5	7424	22	AAK28521	Genomic fragment #
C 34	15.4	62.5	133719	21	AAK64754	Macaca mulatta rha
C 35	15.4	62.5	349980	22	AAH64966	C glutamicum codin
C 36	14.8	61.7	936	21	AAK42786	Arabidopsis thalia
C 37	14.8	61.7	1134	21	AAK45380	Arabidopsis thalia
C 38	14.8	61.7	1180	20	AAK13531	Enterococcus faeca
C 39	14.8	61.7	1197	21	AAK39350	Arabidopsis thalia
C 40	14.8	61.7	2958	13	AAO25264	C glutamicum-deriv
C 41	14.8	61.7	4529	19	AAV16865	Genomic DNA sequen
C 42	14.6	60.8	318	21	AAK75078	Human ORF633
C 43	14.6	60.8	524	21	AAK35552	Arabidopsis thalia
C 44	14.6	60.8	643	21	AAK8385	Fusarium venenatum
C 45	14.6	60.8	828	22	AAH67608	C glutamicum codin

ALIGNMENTS

RESULT 1	AAO25266/C	AAO25266 standard; DNA: 2958 BP.
ID	AAO25266;	
XX	AAO25266;	
AC	AAO25266;	
XX	AAO25266;	
DT	27-NOV-1992 (first entry)	
XX	27-NOV-1992 (first entry)	
DE	Corynebacterium glutamicum ORI sequence.	
XX	Corynebacterium glutamicum ORI sequence.	
KW	Origin of replication; plasmid pKM1519; DS gene; ATCC 13058;	
KW	3-deoxy-D-arabinoheptulosonate-7-phosphate synthase; mutant;	
KW	aromatic amino acid synthesis; retinol inhibition; aspartate kinase; ss.	
XX		
OS	Corynebacterium glutamicum.	
XX	Corynebacterium glutamicum.	
PN	FR2667875-A.	
XX	FR2667875-A.	
PD	17-APR-1992.	
XX	17-APR-1992.	
PF	15-OCT-1991; 91FR-0012701.	
XX	15-OCT-1991; 91FR-0012701.	
PR	15-OCT-1990; 90JP-0273346.	
XX	15-OCT-1990; 90JP-0273346.	
PR	19-JUN-1991; 91JP-0245291.	
XX	19-JUN-1991; 91JP-0245291.	
PA	(AJIN) AJINOMOTO KK.	
XX	(AJIN) AJINOMOTO KK.	
PI	Kojima H, Matsui H, Nakamatsu T, Sato K, Sugimoto M;	
PI	Tanaka A;	
XX	Kojima H, Matsui H, Nakamatsu T, Sato K, Sugimoto M;	
XX	Tanaka A;	
DR	WPI: 1992-185928/23.	
XX	WPI: 1992-185928/23.	
PT	New Corynebacterium plasmids with heat sensitive origin of	
PT	replication - inhibited and expelled at high temp. used to	

PT	transform microorganisms by homologous recombination, esp. for
PT	increasing amino acid yields in fermentation
XX	
PS	Example 1; Page 14; 29pp; French.
CC	The origin of replication from plasmid pHM1519, naturally present in
CC	C. glutamicum, was inserted into shuttle vector pHM4. The resulting
CC	plasmid, pHK4, can replicate in E. coli and Corynebacterium. pHK4 was
CC	treated with hydroxylamine and the products were used to transform
CC	Brevibacterium lactofermentum FERM P-7559. Three kanamycin-resistant
CC	strains were unable to develop at high temperatures. These strains
CC	contained the plasmids pHSC4, pHSC22 and pHSC23, respectively.
CC	See AAQ25263-5.
XX	
SQ	Sequence 2958 BP; 700 A; 854 C; 835 G; 569 T; 0 other;
QY	
DB	Query Match 5 cgaacgttcgttaacgtt 22
	432 CGAACGTCGTTAAAGTT 415
DB	Best Local Similarity 94.4%; Pred. No. 49; Length 2958;
	Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
RESULT 2	
AAQ25263/C	
AAQ25263	standard; DNA: 2958 BP.
AC	AAQ25263;
XX	
DT	27-NOV-1992 (first entry)
XX	
DE	C. glutamicum-derived temp-sensitive ORI present in pHSC4.
XX	
KW	Origin of replication; plasmid pHSC4, DS gene; ATCC 13058;
KW	3-deoxy-D-arabinoheptulosonate-7-phosphate synthase; mutant;
XX	aromatic amino acid synthesis; retroinhibition; aspartate kinase; ss.
OS	Corynebacterium glutamicum.
XX	
FH	Key
FT	mutation
FT	Location/Qualifiers
FT	1543
FT	/*tag- a
FT	/note- "Wild-type G substituted by A"
FT	1546
FT	/*tag- b
FT	/note- "Wild-type G substituted by A"
XX	
FR2667875-A.	
XX	
PD	17-APR-1992.
XX	
PF	15-OCT-1991; 91FR-0012701.
XX	
PR	15-OCT-1990; 90JP-0273348.
XX	
PR	19-JUN-1991; 91JP-0245291.
XX	
PA	(AJIN) AJINOMOTO KK.
PI	Kojima H, Matsui H, Nakamatsu T, Satoh K, Sugimoto M;
PI	Tanaka A;
DR	WPI; 1992-185928/23.
XX	
PT	New Corynebacterium plasmids with heat sensitive origin of
PT	replication - inhibited and expelled at high temp. used to
PT	transform microorganisms by homologous recombination, esp. for
XX	increasing amino acid yields in fermentation
XX	
PS	Claim 3; Page 21; 29pp; French.

CC	The origin of replication from plasmid pHM519,naturally present in
CC	C.glutamicum, was inserted into shuttle vector pHM4. The resulting
CC	plasmid, pHK4, can replicate in E.coli and Corynebacterium. pHM4 was
CC	treated with hydroxylamine and the products were used to transform
CC	Brevibacterium lactofermentum FERM P-7559. Three kanamycin-resistant
CC	strains were unable to develop at high temperatures. These strains
CC	contained the plasmids pHSC4, pHSC22 and pHSC23, respectively. The
CC	mutant ORI sequences were determined. They are each capable of
CC	autonomous replication and maintenance in Corynebacterium but
CC	replication is inhibited and the plasmids are expelled from their
CC	hosts when cultured at 31-37 deg.C. The plasmids contg. the
CC	temperature-sensitive sequences are used to produce transformant
CC	Corynebacterium spp. by homologous recombination. A mutant gene
CC	(such as 3-deoxy-D-arabinonephosphonate-7-phosphate synthase or
CC	aspartate kinase mutated to eliminate retroinhibition) is ligated to
CC	the plasmid contg. the mutant ORI. The wild-type gene undergoes
CC	homologous recombination and is expelled from the cell as part of
CC	the plasmid at a temperature of 31-37 deg.C.
CC	See also AAQ25264-6.
SQ	
XX	Sequence 2958 BP; 702 A; 854 C; 833 G; 569 T; 0 other;
OY	5 cgaacgttcgttaacctt 22 Matches 1/; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Db	432 CGAACGTTTCGTAAAGT 415
RESULT 3	
AAQ25265/c	
ID	AAQ25265 standard; DNA; 2958 BP.
XX	
AC	AAO25265:
XX	
DT	27-NOV-1992 (first entry)
XX	
DE	C.glutamicum-derived temp-sensitive ORI present in pHSC23.
XX	
KW	Origin of replication; plasmid pHSC23; DS gene; ATCC 13058; 3-deoxy-D-arabinoheptulosonate-7-phosphate synthase; mutant; aromatic amino acid synthesis; retroinhibition; aspartate kinase; ss.
XX	
OS	Corynebacterium glutamicum.
XX	
FH	Key
FT	mutation
FT	Location/Qualifiers
FT	756
FT	/tag- a
FT	/note= "Wild-type G replaced by A"
FT	1561
FT	mutation
FT	/tag- b
FT	/note= "Wild-type G replaced by A"
FT	1668
FT	mutation
FT	/tag- c
FT	/note= "Wild-type C replaced by T"
FT	1685
FT	mutation
FT	/tag- d
FT	/note= "Wild-type C replaced by T"
XX	
FR2667875-A.	
PN	
PD	17-APR-1992.
PE	15-OCT-1991; 91FR-0012701.
PR	15-OCT-1990; 90JP-0273348.
PR	19-JUN-1991; 91JP-0245291.
RA	(AJIN) AJINOMOTO KK.
XX	

PI Kojima H, Matsui H, Nakamatsu T, Satoh K, Sugimoto M.
 PI Tanaka A;
 DR WPI: 1992-185928/23.
 XX
 PT New *Corynebacterium* plasmids with heat sensitive origin of
 PT replication - inhibited and expelled at high temp. used to
 PT transform microorganisms by homologous recombination, esp. for
 PT increasing amino acid yields in fermentation
 XX
 PS Claim 5; Page 24; 29pp; French.
 CC The origin of replication from plasmid pHM1519, naturally present in
 CC *C. glutamicum*, was inserted into shuttle vector pHR4. The resulting
 CC plasmid, pHR4, can replicate in *E. coli* and *Corynebacterium*. pHR4 was
 CC treated with hydroxylamine and the products were used to transform
 CC *Brevibacterium lactofermentum* FERM P-7559. Three kanamycin-resistant
 CC strains were unable to develop at high temperatures. These strains
 CC contained the plasmids pHSC4, pHSC22 and pHSC23, respectively. The
 CC mutant ORI sequences were determined. They are each capable of
 CC autonomous replication and maintenance in *Corynebacterium* but
 CC replication is inhibited and the plasmids are expelled from their
 CC hosts when cultured at 31-37 deg.C. The plasmids contg. the
 CC temperature-sensitive sequences are used to produce transformant
 CC *Corynebacterium* spp. by homologous recombination. A mutant gene
 CC (such as 3-deoxy-D-arabinonephosphonate-7-phosphate synthase or
 CC aspartate kinase mutated to eliminate retroinhibition) is ligated to
 CC the plasmid contg. the mutant ORI. The wild-type gene undergoes
 CC homologous recombination and is expelled from the cell as part of
 CC the plasmid at a temperature of 31-37 deg.C.
 CC See AA025263-6.
 XX
 SQ Sequence 2958 BP; 702 A; 851 C; 833 G; 572 T; 0 other;

Query Match 68.3%; Score 16.4; DB 13; Length 2958;
 Best Local Similarity 94.4%; Pred. No. 49;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 cgaacgttcgttaacgtt 22
 |||||
 Db 432 CGACGTCGTTAAAGT 415

RESULT 4
 ID AAX84319/c
 XX AAX84319 standard; DNA: 4142 BP.

AC AAX84319;
 XX
 DT 08-SEP-1999 (first entry)
 XX
 DE Stealth virus nucleic acid clone, SEQ ID NO: 11.
 XX
 KW Stealth virus; detection; diagnosis; infection; ss.
 XX
 OS Stealth virus.
 XX
 PN WO9934019-A1.
 XX
 PD 08-JUL-1999.
 XX
 PF 30-DEC-1998; 98WO-US27744.
 XX
 PR 30-DEC-1997; 97US-0001184.
 XX
 PA (MART/) MARTIN W J.
 XX
 PI Martin WJ;
 XX
 DR WPI: 1999-405521/34.
 XX
 PT Novel strains of stealth virus

XX
 PS Claim 19; Page 42-44; 95pp; English.
 XX
 CC This sequence represents a Stealth virus nucleic acid clone. The
 CC invention relates to a method of detecting and characterising a stealth
 CC virus by reacting a sample suspected of containing a stealth virus with a
 CC probe from a known stealth virus and sequencing the resultant isolated
 CC nucleotide. The method comprises the steps of: (a) isolating DNA or RNA
 CC from a sample suspected of containing a stealth virus, e.g. a culture of
 CC cells showing a viral cytopathic effect; (b) testing the reactivity of
 CC the isolated DNA or RNA with a molecular probe that contains at least 18
 CC or more contiguous nucleotides identical to sequence previously
 CC identified from a stealth virus; and, optionally (c) sequencing the
 CC isolated DNA or RNA molecules that react with the probe. The method is
 CC used to detect stealth virus in a biological product, food or in the
 CC environment. The method is also used to evaluate agents for their
 CC inhibitory or stimulatory effects on stealth virus replication and to
 CC determine capacity of the virus to recombine with and potentially alter
 CC the nucleic acid sequences of a cell or bacterium.
 XX
 SQ Sequence 4142 BP; 1029 A; 964 C; 1082 G; 1065 T; 2 other;

Query Match 66.7%; Score 16; DB 20; Length 4142;
 Best Local Similarity 79.2%; Pred. No. 77;
 Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 tcgtcgaacgttcgttaacgttcg 24
 |||||
 Db 289 TCGTCGAACGCTCTTCACTATCG 266

RESULT 5
 ID AAC47124
 XX AAC47124 standard; DNA: 4414 BP.

AC AAC47124;
 XX
 DT 18-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 52650.
 XX
 KW Hybridisation assay; genetic mapping; gene expression control;
 KW protein identification; signal transduction pathway;
 KW metabolic pathway; promoter; termination sequence; ss.
 XX
 OS Arabidopsis thaliana.

XX
 FN EP1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-0301439.

XX
 PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0130891.
 PR 30-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 04-MAY-1999; 99US-0132407.
 PR 05-MAY-1999; 99US-0132484.
 PR 99US-0132485.

PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135628.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139819.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142350.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155119.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.

PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 66.7%; Score 16; DB 21; Length 4414;
 Best Local Similarity 79.2%; Pred. NO. 76;
 Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 tcgtcgacgttcgttaacgttcg 24
 ||||| | ||||| ||| |||
 Db 230 tcgtccagcttcgttcgacgttcg 253

RESULT 6
 AAF21611/c
 ID AAF21611 standard; DNA: 13732 BP.

AC AAF21611;

DT 04-DEC-2000 (first entry)

XX N. meningitidis partial DNA sequence gnm_32 SEQ ID NO:32.

KW Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
 antigen; vaccine; diagnosis; infection; antibacterial; identification;
 KW Meningococcus B; Menb; ds.

OS Neisseria meningitidis.

PN WO200022430-A2.

PD 20-APR-2000.

PF 08-OCT-1999; 99WO-US23573.

PR 09-OCT-1998; 98US-0103794.

PR 30-APR-1999; 99US-0132068.

PA (CHIR) CHIRON CORP.

PI Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC,
 Masignani V, Galeotti C, Mora M, Ratti G, Scarcelli M, Scarlato V,
 Rapunoli R, Pizzo M.

DR WPI: 2000-318079/27.

XX Isolated nucleotide sequences of Neisseria meningitidis which can be
 PT used in the diagnosis and treatment of N. meningitidis infection and
 PT other Neisserial infections, for example, N.gonorrhoea -
 XX
 XX Claim 7; Page 589-593; 1760pp; English.

CC The present invention describes methods of obtaining immunogenic
 CC proteins from Neisseria genomic sequences. AAF21611 to AAF21614
 CC represent specifically claimed Neisseria meningitidis genomic DNA
 CC sequences. AAF21611 to AAF21614 and AAF21615 to AAF21619 represent
 CC Neisseria DNA sequences and their corresponding proteins. AAF21615 to
 CC AAF21619 and AAF21620 to AAF21624 represent PCR primers used in the
 CC isolation of Neisseria meningitidis DNA sequences; and AAF21622 to
 CC AAF21645 represent Neisseria meningitidis MenB polynucleotide ORF
 CC sequences, which are all used in the exemplification of the present

CC invention. The nucleic acid sequences, protein sequences, and antibodies
 CC against them, can be used in the manufacture of a composition. The
 CC composition can be used as a medicament (or in the manufacture of a
 CC medicament) for treating, preventing or diagnosing infection due to
 CC Neisserial bacteria. For example, some of the identified proteins could
 CC be components of vaccines against Meningococcus B; against all serotypes;
 CC and/or against all pathogenic Neisseriae. Identification of sequences
 CC from the bacterium will also facilitate production of biological probes,
 CC particularly organism-specific probes. Attempts to make efficacious
 CC Meningococcus B vaccines have failed mainly due to antigen tolerance.
 CC Multivalent vaccines have also been tried but none have successfully
 CC overcome antigenic variability. The provision of further, complete
 CC sequences may provide an opportunity to identify secreted or surface
 CC exposed proteins that may be presumed targets for the immune system and
 CC which are not antigenically variable or at least more conserved than
 CC other more variable regions.

XX Sequence 13732 BP; 3488 A; 3771 C; 3357 G; 3116 T; 0 other;

Query Match 66.7%; Score 16; DB 21; Length 13732;
 Best Local Similarity 79.2%; Pred. NO. 75;
 Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 tcgtcgacgttcgttaacgttcg 24
 ||||| | ||||| ||| |||
 Db 13293 TCGCCGACATCCGTCGCGTTCG 13270

RESULT 7
 AAF21611/c
 ID AAF21611 standard; DNA: 349980 BP.

AC AAF21611;

DT 13-MAR-2001 (first entry)

XX Neisseria meningitidis B nucleotide sequence SEQ ID NO:112.

KW Neisseria meningitidis; Neisseria gonorrhoeae; immunogenic; vaccine;
 diagnosis; antigen; detection; infection; gene therapy; antibacterial;
 ds.

OS Neisseria meningitidis.

PN WO20006791-A1.

PD 09-NOV-2000.

PF 08-MAR-2000; 2000WO-US05928.

PR 30-APR-1999; 99US-0132068.

PR 08-OCT-1999; 99WO-US23573.

PR 28-FEB-2000; 2000GB-0004695.

PA (CHIR) CHIRON CORP.

PI Pizzo M, Hickey E, Peterson J, Tettelin H, Venter JC, Masignani V,
 Galeotti C, Mora M, Ratti G, Scarcelli M, Scarlato V, Rapunoli R,
 Frazer CM, Grandi G.

DR WPI: 2000-647603/62.

XX Neisseria meningitidis B full length genome sequence and open reading
 PT frames are used to detect, treat and prevent Neisserial infections -
 XX
 XX Claim 7; Appendix A; 692pp; English.

CC The present invention describes the full length genome of
 CC Neisseria meningitidis B (NMB). The sequences in AAF21544 and AAF21607
 CC to AAF21613 represent fragments of the NMB genomic sequence, as the
 CC sequence was too long to go in a record on its own it was split into 8

CC sequences which overlap each other at the beginning and end of each
CC sequence by 49980 bp (i.e. the last 49980 bp of AAF21544 is repeated at
CC the beginning of AAF21607, the last 49980 bp of AAF21607 are repeated at
CC the beginning of AAF21608, and so on). AAF21545 to AAF21588 encode the
CC Neisseria proteins given in AAB58550 to AAB58593, and AAF21589 to
CC AAF21606 represent PCR primers which are used in the exemplification of
CC the present invention. The NMB genome and fragments from it have
CC antibacterial activity, and can be used in vaccines and gene therapy.
CC Neisseria nucleic acids, proteins and/or antibodies which binds to the
CC proteins can be used in compositions for treating or preventing infection
CC due to Neisserial bacteria or as a diagnostic reagent for detecting the
CC presence of Neisserial bacteria or of antibodies raised to Neisserial
CC bacteria. Computers, computer memory, computer storage medium or computer
CC databases can be used in a search to identify open reading frames (ORFs)
CC or coding sequences within the NMB genome. The DNA sequences provide
CC further opportunities to find antigenic or immunogenic proteins which are
CC more effective in vaccines than the outer membrane proteins currently
CC used.

XX
SQ Sequence 349980 BP; 87189 A; 93501 C; 84627 G; 84663 T; 0 other;

Query Match 66.7%; Score 16; DB 21; Length 349980;
Best Local Similarity 79.2%; Pred. No. 71;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 tcgtcgacgttcgttaacgttcg 24
||| ||||| ||||| |||||
Db 145258 TCGGCGAATCCGCTTCGCGCTTCG 145235

RESULT 8
AAV86097
ID AAV86097 standard; cDNA: 506 BP.
XX
AC AAV86097;
XX
DT 27-APR-1999 (first entry)
XX
DE EST clone D310.
XX
KW Expressed sequence tag; secreted protein; haematopoiesis regulator;
KW tissue growth; activin; inhibin; tumour invasion suppressor; EST; human;
KW chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;
KW receptor; ligand; anti-inflammatory; tumour inhibitor; ds.
XX
OS Homo sapiens.
XX
PN MO9845435-A2.
XX
PD 15-OCT-1998.
XX
PF 10-APR-1998; 98WO-US06954.
XX
PR 10-APR-1997; 97US-0835913.
XX
PA (GENY) GENETICS INST INC.
XX
PI Agostino MJ, Jacobs K, Lavallic ER, McCoy JM, Merberg D;
PI Racie LA, Spaulding V, Treacy M;
XX
DR WPI: 1999-070076/06.
XX
PT New polynucleotides encoding human secreted proteins - derived from
PT e.g. human blood, kidney, foetal lung, placenta, testes, brain,
PT ovary, pituitary, retina and colon cDNA libraries
XX
PS Claim 1; Page 119; 633pp; English.
CC This sequence represents an expressed sequence tag (EST), and is a
CC polynucleotide of the invention. The polynucleotides of the invention are
CC all secreted EST sequences isolated from a variety of human tissue
CC sources. The EST sequences and proteins encoded by them are predicted to

CC have useful biological activities which would make them suitable for
CC treating, preventing or ameliorating medical conditions in humans and
CC animals, although no supporting data is given. Suggested activities
CC include nutritional activity, immune stimulating or suppressing activity,
CC haematopoiesis regulating activity, tissue growth activity,
CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, cachectin/tumour invasion suppressor activity, tumour inhibition
CC therapy. The EST sequences are also stated to be useful for gene
CC therapy.

XX
SQ Sequence 506 BP; 135 A; 133 C; 88 G; 136 T; 14 other;

Query Match 65.8%; Score 15.8; DB 20; Length 506;
Best Local Similarity 89.5%; Pred. No. 98;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 cgtcgacgttcgttaacg 20
||| ||||| ||||| ||
Db 340 cgtcgacgttcgttaacg 358

RESULT 9
AAI13134/c
ID AAI13134 standard; DNA: 9619 BP.
XX
AC AAI13134;
XX
DT 19-MAR-1999 (first entry)
XX
DE Enterococcus faecalis genome contig SEQ ID NO:197.
XX
KW Enterococcus faecalis; contig; detection; Enterococcal infection;
KW vaccine; attenuation; computer readable medium; ds.
XX
OS Enterococcus faecalis.
XX
PN MO9850555-A2.
XX
PD 12-NOV-1998.
XX
PF 04-MAY-1998; 98WO-US08985.
XX
PR 14-NOV-1997; 97US-0066009.
PR 06-MAY-1997; 97US-0044031.
PR 16-MAY-1997; 97US-0046535.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Barash SC, Dillon PJ, Kunsch CA;
XX
DR WPI: 1999-045171/04.
XX
PT New isolated Enterococcus faecalis polynucleotides and polypeptides
PT - used to develop products for the detection of Enterococcus and for
PT use in vaccines for prevention or attenuation of Enterococcus
PT infection.
XX
PS Claim 1; Page 1046-1051; 2084pp; English.
XX
CC A computer readable medium has been developed which has recorded on it
CC 982 nucleotide sequences isolated from the Enterococcus faecalis genome.
CC AAI12938 to AAI13919 represent these nucleotide sequences which are
CC primary nucleotide sequences, also known as contigs. The computer-based
CC system can identify fragments of the Enterococcus faecalis genome with
CC commercial importance. The products can be used to detect the presence
CC of Enterococcus faecalis in samples. They can also be used for
CC diagnosing Enterococcal infection in an animal and monitoring
CC progression of disease, and for identifying agents which can be used to
CC modulate the growth or pathogenicity of Enterococcus faecalis, or
CC another related organism, in vivo or in vitro. In particular the
CC polypeptides encoded by the Enterococcus faecalis nucleotide sequences

CC can be used in vaccines to prevent or attenuate an Enterococcal
 CC infection.
 XX
 SQ Sequence 9619 BP; 2909 A; 1604 C; 2033 G; 3062 T; 11 other;

Query Match 65.88; Score 15.8; DB 20; Length 9619;
 Best Local Similarity 89.58; Pred. No. 94;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 YQ 2 cgtcgaacttcgttaacg 20
 ||| ||||| ||||| |||||
 Db 9133 CGTGAACGTTCTTTAACG 9115

RESULT 10
 AAD03845/C
 ID AAD03845 standard; DNA; 555 BP.

AC AAD03845;
 XX
 DT 31-JUL-2001 (first entry)
 XX

DE Complement of human B cell maturation protein (BCMA) DNA.

KW Human; A Proliferation Inducing Ligand Receptor; APRIL-R; cytostatic;
 KW gene therapy; cancer; nephrotropic; renal disorder; autoimmune disease;
 KW carcinoma; lung; colon; breast; prostate; Grave's disease; hypertension;
 KW systemic lupus erythematosus; SLE; inflammation; cardiovascular disease;
 KW B-cell lympho-proliferative disorder; BCM; immunosuppressive disease;
 KW organ transplantation; HIV; human immunodeficiency virus; TNF;
 KW tumour necrosis factor; BCMA; B cell maturation protein; ds.

XX Homo sapiens.

OS WO200124811-A1.

PN 12-APR-2001.

PD 05-OCT-2000; 2000WO-US27579.

PF 06-OCT-1999; 99US-0157933.

PR 11-FEB-2000; 2000US-0181807.

PR 30-JUN-2000; 2000US-0215688.

XX (BIOJ) BIOGEN INC.

PA (ABOT-) APOTEC R & D SA.

XX

PI Schneider P, Thompson J, Cachero T, Ambrose C, Rennett P;

XX WPI: 2001-266242/27.

DR Treating a mammal for a condition associated with undesired cell

XX proliferation such as cancer or carcinoma, comprises administering a

PT composition comprising A proliferation Inducing Ligand Receptor

PT (APRIL-R) antagonist -

XX Disclosure; Fig 3A; 85pp; English.

XX The invention relates to a method of treating a mammal for a condition

CC associated with undesired cell proliferation such as cancer or

CC carcinoma. The method involves administering a composition comprising

CC B cell maturation protein (BCM or BCMA) antagonist that antagonises the

CC interaction between APRIL and its cognate receptor(s). This method is

CC useful for treating undesired cell proliferation such as cancer or

CC carcinoma e.g. human lung carcinoma, colon carcinoma, breast carcinoma,

CC prostate carcinoma, and other carcinomas whose proliferation is modulated

CC by APRIL. It is also useful for treating autoimmune diseases (Grave's

CC disease, systemic lupus erythematosus-SLE); hypertension, cardiovascular

CC diseases, renal disorders, B-cell lympho-proliferative disorders,

CC immunosuppressive diseases, organ transplantation, inflammation and

CC human immunodeficiency virus (HIV), and for treating, suppressing or

CC altering an immune response involving a signalling pathway between

CC APRIL-R and its ligand. APRIL-R DNA is also useful in gene therapy.

CC The present sequence is a complement of human APRIL-R also referred

CC as BCMA DNA.

XX SQ Sequence 555 BP; 152 A; 132 C; 115 G; 156 T; 0 other;

Query Match 65.08; Score 15.6; DB 22; Length 555;
 Best Local Similarity 81.88; Pred. No. 1.2e+02;
 Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 YQ 1 tcgtcgaacttcgttaacgtt 22
 ||||| ||| ||| |||||
 Db 513 TCGTCGACCGTCGGAACGTT 492

RESULT 11
 AAV75360
 ID AAV75360 standard; DNA; 608 BP.

AC AAV75360;

XX 16-MAR-1999 (first entry)

DE Staphylococcus aureus contig SEQ ID #1049.

XX Computer readable medium; vaccine; S.aureus infection; immunodetection;

KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;

KW skin infection; surgical wound infection; scalded skin syndrome;

KW toxic shock syndrome; ds.

XX Staphylococcus aureus.

OS EP786519-A2.

PN 30-JUL-1997.

PD 07-JAN-1997; 97EP-0100117.

PF 05-JAN-1996; 96US-0009861.

PR (HUMA-) HUMAN GENOME SCI INC.

XX Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA;

PI Rosen CA;

XX WPI: 1997-374922/35.

DR Polynucleotide(s) and proteins derived from Staphylococcus aureus

XX stored on computer readable medium and used in the production of

PT anti-S.aureus vaccines

XX Claim 1; Page 1809; 3271pp; English.

XX This sequence represents one of 5191 Staphylococcus aureus DNA sequences

CC of the invention. The DNA sequences are recorded on a computer readable

CC medium, preferably selected from a floppy or hard disk, random access

CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using

CC the S.aureus DNA sequences allows putative functions to be assigned so

CC that protein-encoding or regulatory regions of commercial, therapeutic or

CC industrial importance can be obtained. Specifically, sequences which are

CC likely to encode antigens have been identified and these polypeptides can

CC be used in a vaccine composition against S.aureus infection. The

CC polypeptides can also be used in a kit for the immunodetection of

CC S.aureus in a sample. S.aureus is implicated in numerous human diseases,

CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,

CC skin and surgical wound infections, scalded skin syndrome, toxic shock

CC syndrome, etc. Organisms transformed with the DNA sequences can be used

CC for recombinant production of the polypeptides. The new DNA sequences

CC (and their fragments) are useful as primers or probes for isolating

CC homologues of any of the S.aureus DNA sequences contained on the

CC computer readable medium.

Tue Mar 12 10:44:34 2002

us-09-802-518-9.rng

Page 10

Search completed: March 11, 2002, 20:59:57
Job time: 4180 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2002, 20:55:45 ; Search time 1535.11 Seconds
(without alignments)
257.918 Million cell updates/sec

Title: US-09-802-518-9
Perfect score: 24
Sequence: 1 tcgtcgaacttcgttaacttcgc 24

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenDbml:*

- 1: gb_da:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vl:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_om:*
- 20: em_or:*
- 21: em_ov:*
- 22: em_pat:*
- 23: em_ph:*
- 24: em_pl:*
- 25: em_ro:*
- 26: em_sts:*
- 27: em_sy:*
- 28: em_un:*
- 29: em_vl:*
- 30: em_htgo_hum:*
- 31: em_htgo_inv:*
- 32: em_htgo_rod:*
- 33: em_htg_hum:*
- 34: em_htg_inv:*
- 35: em_htg_rod:*
- 36: em_htg_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	18.4	76.7	148263	8	AC051624	AC051624 Genomic S
2	18.2	75.8	1063	1	AB011203	AB011203 Streptococ
3	18.2	75.8	166050	2	AC079766	AC079766 Homo sapi
4	17.6	73.3	1680	1	HM08FA	X73823 H mediterr
5	17.6	73.3	43075	7	U88874	U88874 Streptococ
6	17.6	73.3	169696	2	AC083943	AC083943 Oryza sat
7	17.4	72.5	121524	8	AF041468	AF041468 Guillard
8	17.2	71.7	5905	5	AF007218	AF007218 Tetradon
9	17.2	71.7	183317	2	CNS07ERT	AL513003 Oryza sat
10	17	70.8	224	8	AY024037	AY024037 Oryza sat
11	17	70.8	101803	2	AP003370	AP003370 Oryza sat
12	17	70.8	104718	2	AP003987	AP003987 Oryza sat
13	17	70.8	149498	2	AP003196	AP003196 Oryza sat
14	17	70.8	150379	8	AP003074	AP003074 Oryza sat
15	17	70.8	150641	2	AP003211	AP003211 Oryza sat
16	16.8	70.0	81580	8	AP000736	AP000736 Arabidops
17	16.8	70.0	98017	8	AC027033	AC027033 Arabidops
18	16.8	70.0	136047	36	AC069470	AC069470 Arabidops
19	16.8	70.0	140019	2	AC093178	AC093178 Oryza sat
20	16.8	70.0	347950	1	AP003013	AP003013 Mesorhizo
21	16.6	69.2	645	3	SMU35432	U35432 Schistosoma
22	16.6	69.2	766	3	ASNS09	AF130132 Anopheles
23	16.6	69.2	3414	3	DM021123	U21123 Drosophila
24	16.6	69.2	9180	1	AE005246	AE005246 Escherich
25	16.6	69.2	10294	1	AE001700	AE001700 Thermotog
26	16.6	69.2	28224	2	AC014904	AC014904 Drosophill
27	16.6	69.2	39708	3	CEP35G12	U46242 Caenorhabdi
28	16.6	69.2	70283	2	AC092870	AC092870 Takifugu
29	16.6	69.2	72019	2	AC009914	AC009914 Drosophill
30	16.6	69.2	103683	2	AC012387	AC012387 Drosophill
31	16.6	69.2	106559	2	AC020331	AC020331 Drosophill
32	16.6	69.2	107967	9	AL333701	AL333701 Human DNA
33	16.6	69.2	128444	2	AC019924	AC019924 Drosophill
34	16.6	69.2	161172	3	AC007839	AC007839 Drosophill
35	16.6	69.2	178692	8	AC024594	AC024594 Oryza sat
36	16.6	69.2	180495	3	AC009847	AC009847 Drosophill
37	16.6	69.2	188633	3	AC007175	AC007175 Drosophill
38	16.6	69.2	201400	2	AC067780	AC067780 Homo sapi
39	16.6	69.2	229896	14	AF232689	AF232689 Rat cytom
40	16.6	69.2	262278	1	AP002552	AP002552 Escherich
41	16.6	69.2	287859	3	AE003797	AE003797 Drosophill
42	16.6	69.2	296202	3	AE003785	AE003785 Drosophill
43	16.6	69.2	302547	3	AE003453	AE003453 Drosophill
44	16.4	68.3	2958	6	E04235	E04235 Temperature
45	16.4	68.3	2958	6	E04236	E04236 Temperature

ALIGNMENTS

RESULT 1
AC051624/c
LOCUS
DEFINITION
Genomic Sequence for Oryza sativa, Nipponbare strain, clone
OSJNB0036B06, from Chromosome 10, complete sequence.
AC051624
VERSION
AC051624.6 GI:14165313
KEYWORDS
HTG.
ORGANISM
Oryza sativa.
SOURCE
Oryza sativa
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
Spermatophyta: Magnoliophyta: Liliopsida: Poales: Poaceae:
Ehrhartoideae: Oryzaceae: Oryza.
(bases 1 to 148263)

REFERENCE
AUTHORS
Du, H., Minx, P., Abbott, A., Doeber, A., de la Bastide, M.,
Spiegel, L., Nascimben, L., Preston, R., Kirchoff, K., King, L.,
Vill, M.D., Baker, J., Zutter, T., Santos, L., Bell, M., Miller, B.,
Kuit, K., Rodriguez, S., Cunnius, D.M., Baliga, V., Shah, R., Bahrel, A.,
O'Shaughnessy, A., Palmer, L., Yang, C., Dethia, N. and McCombie, W.R.
Genomic Sequence for Oryza sativa, Nipponbare strain, clone

TITLE

JOURNAL OSCNB0036B06, from Chromosome 10, complete sequence
 REFERENCE 2 (bases 1 to 148263)
 AUTHORS MComble, W.R.
 JOURNAL Direct Submission
 Submitted (15-APR-2000) Uta Annenberg Hazen Genome Center, Cold Spring Harbor Laboratories, 1, Bungtown Road, Cold Spring Harbor, NY 11724, USA
 REFERENCE 3 (bases 1 to 148263)
 AUTHORS MComble, W.R.
 JOURNAL Direct Submission
 Submitted (22-MAY-2001) Uta Annenberg Hazen Genome Center, Cold Spring Harbor Laboratories, 1, Bungtown Road, Cold Spring Harbor, NY 11724, USA
 COMMENT On May 22, 2001 this sequence version replaced g1:9972290. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

FEATURES
 source
 1..148263
 /organism="Oryza sativa"
 /strain="Nipponbare"
 /db_xref="taxon:4530"
 /chromosome="10"
 /clone="OSNB0036B06"
 /clone_1lb="Ecot1"
 /note="The sequence contains a dinucleotide (TA) repeat from base position 37007 to 37090 in which the exact length is unknown. Assembly of the database is consistent with PCR."

misc_feature
 /note="The sequence contains a dinucleotide (TA) repeat from base position 37007 to 37090 in which the exact length is unknown. Assembly of the database is consistent with PCR."

BASE COUNT 36837 a 35840 c 36643 g 38943 t
 ORIGIN

Query Match 76.7%; Score 18.4; DB 8; Length 148263;
 Best Local Similarity 95.0%; Pred. No. 74;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 cgaagtcgttaacgtcg 24
 ||| ||||| ||||| |||||
 Db 65607 CGAGCTGCTTAACGTCG 65588

RESULT 2
 LOCUS AB011203 1063 bp DNA BCT 27-JUL-2000
 DEFINITION Streptococcus pneumoniae pbp2x gene for penicillin binding protein 2X, partial cds, isolate #13/S04.
 ACCESSION AB011203
 VERSION AB011203.1 GI:4519225
 KEYWORDS pbp2x; penicillin binding protein 2X.
 SOURCE Streptococcus pneumoniae (isolate:#13/S04) DNA.
 ORGANISM Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae; Streptococcus.
 REFERENCE 1 (sites)
 AUTHORS Asahi, Y., Takeuchi, Y. and Ubukata, K.
 TITLE Diversity of substitutions within or adjacent to conserved amino acid motifs of penicillin-binding protein 2X in cephalosporin-resistant Streptococcus pneumoniae isolates
 JOURNAL Antimicrob. Agents Chemother. 43 (5), 1252-1255 (1999)
 MEDLINE 99240374
 REFERENCE 2 (bases 1 to 1063)
 AUTHORS Asahi, Y. and Ubukata, K.
 TITLE Direct Submission
 JOURNAL Submitted (13-FEB-1998) to the DDBJ/EMBL/GenBank databases. Kimiko Ubukata, Meiji Seika Kaisha, LTD, Pharmaceutical Research Center; 760 Morooka-cho, Kohoku-ku, Yokohama, Kanagawa 222-8567, Japan

(E-mail: Kimiko_Ubukata@meiji.co.jp, Tel: 81-45-545-3106, Fax: 81-45-545-3129)
 LOCATION/Qualifiers
 1..1063
 /organism="Streptococcus pneumoniae"
 /isolate="#13/S04"
 /db_xref="taxon:1313"
 1..1063
 /gene="pbp2x"
 <1..>1063
 /gene="pbp2x"
 /note="corresponding to nucleotides 1018-2080 and codon 256-693 in the pbp2x: EMBL Acc#No. X16367"
 /codon_start=1
 /transl_table=1
 /product="penicillin binding protein 2X"
 /protein_id="BAA75526.1"
 /db_xref="GI:4519226"
 /translation="VDEKDYVTLSPLOSPMETQDAFOEKYKGYMTATLYSAKTG
 EIIATQRPFNADTKDKIDFVNRDILQSNYEPSTKVMMLAAIDNKPEGGE
 YFNSEELKADATIRMDVNEGLSGTMTFSGFAHSSNGMTLLBOKGDATWLDY
 LNRKRGVPRRGLTDEYTGOLPADNTVNIAMSAFGGISTVTOIMRAFTAIANDGV
 MLERKTSATKIDPDOSVRSKREYVGNPYSKEASTRDMHWVGTDPYGTNRNS
 TGRATVNPQONVALKSGTAQIADERNKGLTGSTNTFSVSMHPANPDPIILVTV
 QDEHYSGIQLGEFANPILF"
 244..255
 /gene="pbp2x"
 /note="STMK homology box"
 244..246
 /gene="pbp2x"
 /note="active site serin; unclassified site"
 418..426
 /gene="pbp2x"
 /note="SSN homology box"
 874..882
 /gene="pbp2x"
 /note="KSG homology box"
 212 c 249 g 299 t

BASE COUNT 303 a 212 c 249 g 299 t
 ORIGIN

Query Match 75.8%; Score 18.2; DB 1; Length 1063;
 Best Local Similarity 87.0%; Pred. No. 89;
 Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 tcgtcgaagtcgttaacgtc 23
 ||||| ||| ||||| |||||
 Db 379 TCGTCAACCTGCTTAACGTCG 357

RESULT 3
 LOCUS AC079766 166050 bp DNA HTG 08-MAY-2001
 DEFINITION Homo sapiens chromosome RPc1-11 clone RP11-188P17, WORKING DRAFT SEQUENCE, 31 unordered pieces.
 ACCESSION AC079766
 VERSION AC079766.4 GI:13775303
 KEYWORDS HTG; HTGS-PHASE1; HTGS-DRAFT.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 166050)
 AUTHORS Waterston, R.H.
 TITLE The sequence of Homo sapiens clone
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 166050)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (10-SEP-2000) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 COMMENT On Apr 24, 2001 this sequence version replaced g1:12718914.

```

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H_NH0188P17
----- Summary Statistics -----
Sequencing vector: MJ3; 68%
Sequencing vector: plasmid; 9%
Chemistry: Dye-primer ET; 68% of reads
Chemistry: Dye-terminator Big Dye; 9% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 143207 bases at least Q40
Consensus quality: 149551 bases at least Q30
Consensus quality: 152965 bases at least Q20
Insert size: 167000; agarose-fp
Insert size: 163050; sum-of-contigs
Quality coverage: 3.54 in Q20 bases; agarose-fp
Quality coverage: 3.81 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 31 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1 1750: contig of 1750 bp in length
* 1751 1850: gap of unknown length
* 1851 3417: contig of 1567 bp in length
* 3418 3517: gap of unknown length
* 3518 4872: contig of 1355 bp in length
* 4873 4972: gap of unknown length
* 4973 7244: contig of 2272 bp in length
* 7245 7345: gap of unknown length
* 7345 8892: contig of 1548 bp in length
* 8893 8992: gap of unknown length
* 8993 12042: contig of 3050 bp in length
* 12043 12142: gap of unknown length
* 12143 14327: contig of 2185 bp in length
* 14328 14427: gap of unknown length
* 14428 16393: contig of 1966 bp in length
* 16394 16493: gap of unknown length
* 16494 18729: contig of 2236 bp in length
* 18730 18829: gap of unknown length
* 18830 21162: contig of 2333 bp in length
* 21163 21262: gap of unknown length
* 21263 23860: contig of 2598 bp in length
* 23861 23960: gap of unknown length
* 23961 26386: contig of 2426 bp in length
* 26387 26487: gap of unknown length
* 26487 30147: contig of 3661 bp in length
* 30148 30247: gap of unknown length
* 30248 34442: contig of 4195 bp in length
* 34443 34542: gap of unknown length
* 34543 37265: contig of 2723 bp in length
* 37266 37365: gap of unknown length
* 37366 40885: contig of 3520 bp in length
* 40886 40985: gap of unknown length
* 40986 44345: contig of 3360 bp in length
* 44346 44445: gap of unknown length
* 44446 49014: contig of 4569 bp in length
* 49015 49114: gap of unknown length
* 49115 53415: contig of 4301 bp in length
* 53416 53515: gap of unknown length
* 53516 57695: contig of 4180 bp in length
* 57696 57795: gap of unknown length
* 57796 63198: contig of 5403 bp in length
* 63199 63298: gap of unknown length
* 63299 70267: contig of 6969 bp in length
* 70268 70367: gap of unknown length

```

```

FEATURES
      .source
* 70368 75697: contig of 5330 bp in length
* 75698 75797: gap of unknown length
* 75798 83050: contig of 7253 bp in length
* 83051 83150: gap of unknown length
* 83151 92487: contig of 9337 bp in length
* 92488 92587: gap of unknown length
* 92588 101334: contig of 8747 bp in length
* 101335 101434: gap of unknown length
* 101435 110619: contig of 9185 bp in length
* 110620 110719: gap of unknown length
* 110720 120293: contig of 9574 bp in length
* 120294 120393: gap of unknown length
* 120394 132322: contig of 11929 bp in length
* 132323 132422: gap of unknown length
* 132423 146742: contig of 14320 bp in length
* 146743 146842: gap of unknown length
* 146843 166050: contig of 19208 bp in length.
      Location/Qualifiers
1. 166050
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="RPC1-11"
/clone="RP11-188P17"
1. 1750
/note="assembly_name:Contig22"
1851..3417
/note="assembly_name:Contig23"
3518..4872
/note="assembly_name:Contig24"
4973..7244
/note="assembly_name:Contig25"
7345..8892
/note="assembly_name:Contig26"
8993..12042
/note="assembly_name:Contig27"
12143..14327
/note="assembly_name:Contig28"
14428..16393
/note="assembly_name:Contig29"
16494..18729
/note="assembly_name:Contig30"
18830..21162
/note="assembly_name:Contig31"
21263..23860
/note="assembly_name:Contig32"
23961..26386
/note="assembly_name:Contig33"
26487..30147
/note="assembly_name:Contig34"
30248..34442
/note="assembly_name:Contig35"
34543..37265
/note="assembly_name:Contig36"
37366..40885
/note="assembly_name:Contig37"
40986..44345
/note="assembly_name:Contig38"
44446..49014
/note="assembly_name:Contig39"
49115..53415
/note="assembly_name:Contig40"
53516..57695
/note="assembly_name:Contig41"
57796..63198
/note="assembly_name:Contig42"
63299..70267
/note="assembly_name:Contig43"
70368..75697
/note="assembly_name:Contig44"
75798..83050
/note="assembly_name:Contig45"
83151..92487
/note="assembly_name:Contig46"

```


CDS

/codon_start=1
/transl_table=11
/product="ORF2"
/protein_id="AAC79518.1"
/db_xref="GI:2444082"
/translation="MKMDVNRVPTLYLKEFTFGGLITGFIPEPFWLMEVGA
LFLRALFMETIKRPMTEKOEKOLKIERAKKEFRSRNPEPKMAEARAVCPCKST
DVEFVQOKRSISGRKAAAGTIMGVVALAGFACKKCKKEHCNCNAVTTTK"
complement(1706..2092)
/codon_start=1
/transl_table=11
/product="ORF3"
/protein_id="AAC79519.1"
/db_xref="GI:2444083"
/translation="MKIDELKEVYISLFIPELMEWSSGFEPDLRIIVYNSLSKQ
EHEKYLIEHGVNHDPRKHYKRLLOVEQADRFMRILRYDLKSTDIYDFNWRFA
AQNYSITGEAMIDDEFKIQOOSYI"
complement(2082..2486)
/note="putative repressor"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="ORF4"
/protein_id="AAC79520.1"
/db_xref="GI:2444084"
/translation="MKGDYIMKENTAOQLAQMKEFNRLROYDILFQSEKFOELGTL
GKSALSOYVSGSIFPDODKVLILSTLGVSEAMLMGYDTNESTDVKKESIDSLNREK
VYMFQKPLSDVDYKRIEQIIRLSIEVGEDD"
2635..2874
/codon_start=1
/transl_table=11
/product="ORF5"
/protein_id="AAC79521.1"
/db_xref="GI:2444085"
/translation="MIKYDYSKLGKIYEVGTGKKFAQSNMLSEKSIASATNNKRC
WQOSEIKKACELQIPSDIEINSYPPRYKQDIEHN"
2888..3169
/codon_start=1
/transl_table=11
/product="ORF6"
/protein_id="AAC79522.1"
/db_xref="GI:2444086"
/translation="MNPDKPLADRFNLTLYVADKSKAPDIDOTPLILAKQCE
MAGIGYTFELRTNLDGFPKIDKRGSOVRIPRQVROWENNMODIA"
3330..3358
/codon_start=1
/transl_table=11
/product="ORF7"
/protein_id="AAC79523.1"
/db_xref="GI:2444087"
/translation="MKNLFKWLISDEKBOKEPWTPEYENRKYEBRHKREQLLMK"
3616..4089
/codon_start=1
/transl_table=11
/product="ORF8"
/protein_id="AAC79524.1"
/db_xref="GI:2444088"
/translation="MATLYELTDQLLEIYNMVDDETKLDTLEAIDMTDYENKVEGY
GVKYLSEADIEARKKEKRLDKNSDOSKIDKLTALAVAMERGOTRVDTLFEKV
GPKSAVYVNEKLLPKYQIATYKPKDKTKLELLSKSHIGATLEBRNINIR"
4086..4787
/note="contains a putative ATP/GTP binding site between
amino acids 21-28"
/codon_start=1
/transl_table=11
/product="ORF9"
/protein_id="AAC79525.1"
/db_xref="GI:2444089"
/translation="MKIIRANDIORTKMRILLYKAGLGKTSLIKNNPQKTLVSLD
NSSKVLAGEVNDIIDREHPEFTETFEITLQENYKENVLINDINSRPSQDVLFE
OGKRSNGNSNELQHTSQMTNFFLRFLITYISKPIINIYTANEDTHELLETGQILTO
VYQIILASVANDLGLTDVGRIVNAKTGAGGLIEGSEGYANRNRLRTACKIED
LRFKGDSTKEIPE"

CDS

4762..6093
/note="putative helicase"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="ORF10"
/protein_id="AAC79526.1"
/db_xref="GI:2444090"
/translation="MELRNYSQSLVNDIKOSILRGNKRIWQSPRSKGTVAHAHAK
GATDKNTLLEFSHRKEINBOVNTFERNODMNLVYIDSVTKIARNDLQPSIIL
IDEAHVYKATYLLKIEYNSIYLMETGTPARLDGSGPDIDDIYLGSKWMLDN
GNIAPEKYVAPSLIDTNTLKRGGGEFLKRSVDDPMKRVYIGDYIRHYEKLRKQALV
YTHSEVASESVSNFNEOGYSTIAISGKTPEYEDRAMQAFRQKTLITMNCLEFREG
IDENVDVCIIMLRPTQSLIYQFAMALNPRBKTAIITIDHGVNDRHGLPNDRREM
SIRGVNNTKRLMLKEPTPTPCOCYATWASARIICWGCHEKPKRKEVEIIRIEL
EERROEVAASVEFEVTSDDQATVEELKEPAKHOGYKGMVYIOOKEKNIR"
6100..6555
/note="function unknown"
/codon_start=1
/transl_table=11
/product="ORF11"
/protein_id="AAC79527.1"
/db_xref="GI:2444091"
/translation="MFTIDYQAKFEFSIKDGYETVIDLAKODATOGADYDIPRR
IRKFOEPONNITFYRIIPAKKEDKGPVASSIMNIAAAGIPGTRKSSLEDLNQLE
GKALATYVKNKESMOCQKTYENLVKRLVTDIPLPEVNTIEISIDLPE"
6564..7373
/codon_start=1
/transl_table=11
/product="ORF12"
/protein_id="AAC79528.1"
/db_xref="GI:2444092"
/translation="MVDYAINYQMGYSVIPISKNGKTPILSPADKRPMTENDIRRWY
RDNDPANIALTDTFEPVYIDVMHGDVGLTNLRMMHARLIPPTLQATPSSGRIYLL
KLIDNPHISQNIIEGVDIKAYHANNIIVPSSNKGYYEMDTVHSPKGSITAEADIL
ALIVLQMKREPIYSYVSSPFSADSVSTKTKTLFESILIGFGGGRNNALAEVQGG
LLRGVDPEITYHLAKKANNNTQEPIDDKEREFKSMLEIKRIGIND"
7360..8874
/note="putative primase"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="ORF13"
/protein_id="AAC79529.1"
/db_xref="GI:2444093"
/translation="MTTIDPDYREOFASSLSPEKPSREGIKNKLKAYRNDMEFERE
KEENDDKEKPALPELAVAKGLNTYVITLENGKVAIYDPERGYOKDRYAYOLIY
VLEPTFNETKCRNVLFIILSNKREYENNMWDEPEYRDYRRILYKNGIYDKRRKK
LLSFIDYKFINSTIETELVENAKPPTINGMDVDSMLDMSGSELYELLMOVYAAASL
NGNSYKRSIWLVGNGNDGKTFQOLLISNVLGNVAPLAKINQSEKFGALIEGKTV
ITIGDVOAGIYDSSNSVSTGEPVPSIEKKSPTPLAOFKKNVIOSTANMAPSEKATN
SNGTYRRIVILPEKRTGINDNNAIDDDYINREYVLEYVLMKAINIDFSEPKATN
OERQGEFEENNTYKFLNELYSDVSRVPRVLEMYVYRMECHGNTIDPKSNFKEK
ELAQNLPYGVWKDKOKPLDFPNPTKDKPYVWHDPNFWMDENAKKAVVAVVYQ"
9128..9553
/codon_start=1
/transl_table=11
/product="ORF14"
/protein_id="AAC79530.1"
/db_xref="GI:2444094"
/translation="METIYLEINQHEHPENENISWYIDFSEATYIDVQCNALIVHFE
KVGDYVDEYDYGKIVVAFIYEDNKAETPKKMKMLEFEDKIDKTEKTLFIDDPFDNN
WATKGDEYVDTFNNCLFQOASMAVYTDQIVSIEWLQNH"
9601..9861
/codon_start=1
/transl_table=11
/product="ORF15"
/protein_id="AAC79531.1"
/db_xref="GI:2444095"
/translation="MVRIRNGKYRMADGRWPTDCAKRGCDLFGFRPDQIIFTEYK
NEKGLRDQCKKEAMKRGALVGVARSYKADADIVNEKRC"
9845..10081
/codon_start=1

CDS

```

/transl_table=11
/product="ORE16"
/protein_id="AAC79532.1"
/db_xref="GI:2444096"
/translation="NRKNYKDKLVGIYAPGHYEHSVLEETQFSRMPWDHEDNEEF
SKLSISTKRLKRLITLEQLPDELLARNGELCK"
10072..10245

```

Query Match	73.3%	Score 17.6;	DB 7;	Length 43075;
Best Local Similarity	83.3%	Pred. No. 1.9e+02;		
Matches 20;	Conservative 0;	Mismatches 4;	Indels 0;	Gaps 0;

QY	1	tcgtcgaa	cggttcg	taacgttcg	24
Db	28112	TCGTGGA	CCCTTCG	TATCGATAG	28085

RESULT	6
AC083943	
LOCUS	
DEFINITION	AC083943 169696 bp DNA HTG 25-JUL-2001
	Oryza sativa clone OSJNBa0044A10, WORKING DRAFT SEQUENCE, 4
	unordered pieces.
ACCESSION	AC083943
VERSION	AC083943.6 GI:15011678
KEYWORDS	HTG; HTGS-PHASE1; HTGS-DRAFT.
SOURCE	Oryza sativa.
ORGANISM	Oryza sativa

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

1 (bases 1 to 169696)
McCombie, W. R.
Rice genomic sequence
Unpublished
2 (bases 1 to 169696)
McCombie, W. R.
Direct Submission
Submitted (07-OCT-2000) Lita Annenberg Hazen Genome Center, Cold

COMMENT

Center: Lita Annenberg Hazen Genome Center, Cold Spring Harbor
Laboratory
Center code: CSHL
Web site: <http://www.cshl.org/genseq>
Contact: mccombbie@cshl.org
----- Project Information
Center project name: OSUNBA0044A10
Center clone name: OSUNBA0044A10

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

*	1	311206:	contig of 131206 bp in length
*		131297:	gap of unknown length
*		151392:	contig of 27730 bp in length
*		151912:	gap of unknown length
*		159309:	contig of 6776 bp in length
*		160088:	gap of unknown length
*		166265:	gap of unknown length
*		166280:	contig of 3427 bp in length.
*		169696:	

```

FEATURES
source      Location/Qualifiers
1. 169696   /organism="Oryza sativa"
           /db_xref="taxon:4530"

```

```

/clone="OSJNBa0044A10"
BASE COUNT 47992 a 36479 c 35957 g 48703 t 565 others
ORIGIN

```

Query Match	73.3%;	Score 17.6;	DB 2;	Length 169696;
Best Local Similarity	83.3%;	Pred. No. 1.9e+02;		
Matches 20;	Conservative 0;	Mismatches 4;	Indels 0;	Gaps 0;

Qy 1 tcgtcgaa cgttcg ttaacgttcg 24
||| | ||||| |||||
Db 17018 TCGTTAGCGTTCGTTACAGTTCG 17041

[illegible]

ORGANISM	REFERENCE	AUTHORS	TITLE
Chloroplast <i>Galliardia theta</i>			
Eukaryote: Cryptophyta: Cryptomonadaceae; <i>Galliardia</i>			
1 (bases 47701 to 48415)			
Douglas, S. E. and Durrford, D. G.			
The small subunit of ribulose-1,5-bisphosphate carboxylase			

plastid-encoded in the chlorophyll *c*-containing alga *Cryptomonas*
ph
JOURNAL Plant Mol. Biol. 13 (1), 13-20 (1989)

JOURNAL	Plant Mol. Biol.	13	(1)	13-20	(1989)
MEDLINE	93357429	2	(bases 45872 to 47981)		
REFERENCE	Douglas, S.E., Durnford, D.G. and Morden, C.W.				
AUTHORS	Nucleotide sequence of the gene for the large subunit of				
TITLE	ribulose-1,5-bisphosphate carboxylase/oxygenase from the				
	chlorophyll <i>c</i> -containing Alga <i>Cryptomonas</i> f: evidence supporting				
JOURNAL	the polyphyletic origin of plastids				
	J. Phycol.	26,	500-508	(1990)	

JOURNAL J. Phycol. 26, 500-508 (1990)
REFERENCE 3 (bases 43739 to 44938)
AUTHORS Douglas, S.E. and Durnford, D.G.
TITLE Nucleotide sequence of the genes for ribosomal protein S4 and tRNA^{Arg} from the chlorophyll *c*-containing alga *Cryptomonas* ph.
JOURNAL Nucleic Acids Res. 18 (7):1903 (1990)

MEDLINE	4 (bases 18535 to 19311)
REFERENCE	90245597
AUTHORS	Douglas, S. E. and Durnford, D. G.
TITLE	Sequence analysis of the plastid rDNA spacer region of the chlorophyll <i>c</i> -containing alga <i>Cryptomonas</i> phl
JOURNAL	DNA. Seq. 1 (1), 55-62 (1990)

MEDLINE	5 (bases 34539 to 35380)
REFERENCE	92119320
AUTHORS	Reith, W. and Douglas, S.
TITLE	Localization of beta-phycoerythrin to the thylakoid lumen of
JOURNAL	Cryptomonas phl does not involve a signal peptide
	Plant Mol. Biol. 15 (4):585-592 (1990)

MEDLINE	91338697
REFERENCE	6 (bases 110917 to 113854)
AUTHORS	Douglas, S.E.
TITLE	Unusual organization of a ribosomal protein operon in the plastid genome of <i>Cryptomonas</i> phl: evolutionary considerations
JOURNAL	Curr. Genet. 19 (4), 289-294 (1991)

MEDLINE	91330343
REFERENCE	7 (bases 40675 to 42376)
AUTHORS	Douglas,S.E. and Turner,S.
TITLE	Molecular evidence for the origin of plastids from a cyanobacterium-like ancestor
JOURNAL	J. Mol. Evol. 33 (3), 267-273 (1991)

MEDLINE	92099311
REFERENCE	8 (bases 96129 to 98906)
AUTHORS	Wang, S.L. and Liu, X.Q.
TITLE	The plastid genome of <i>Cryptomonas phi</i> encodes an hsp70-like protein, a histone-like protein, and an acyl carrier protein


```

JOURNAL      Proc. Natl. Acad. Sci. U.S.A. 88 (23), 10783-10787 (1991)
MEDLINE      92073372
REFERENCE    9 (bases 106789 to 108216)
AUTHORS      Douglas,S.E.
TITLE        A secy homologue is found in the plastid genome of Cryptomonas phi
JOURNAL      FEBS Lett. 298 (1), 93-96 (1992)
MEDLINE      92183838
REFERENCE    10 (bases 42198 to 44153)
AUTHORS      Douglas,S.E. and Keith,M.E.
TITLE        A bchl homolof, encoding a subunit of Mg chelatase, is located on
JOURNAL      the plastid genomes of red and cryptomonad algae
AUTHORS      J. Mar. Biotechnol. 1, 135-141 (1993)
TITLE        11 (bases 82327 to 84479)
JOURNAL      Douglas,S.E. and Murphy,C.A.
AUTHORS      Structural, transcriptional and phylogenetic analyses of the atpB
TITLE        gene cluster from the plastid of Cryptomonas F (Cryptophyceae)
JOURNAL      J. Phycol. 30, 329-340 (1994)
REFERENCE    12 (bases 98901 to 114602)
AUTHORS      Wang,S.L., Liu,X.Q. and Douglas,S.E.
TITLE        The large ribosomal protein gene cluster of a cryptomonad plastid:
JOURNAL      gene organization, sequence and evolutionary implications
AUTHORS      Biochem. Mol. Biol. Int. 41 (5), 1035-1044 (1997)
TITLE        13 (bases 61067 to 68605)
JOURNAL      Leitsch,C.E.W., Kowallik,K.V. and Douglas,S.E.
AUTHORS      The atpA gene cluster of a cryptomonad, Guillardia theta: A piece
TITLE        in the puzzle of chloroplast genome development
JOURNAL      J. Phycol. (1998) In press
REFERENCE    14 (bases 1 to 121524)
AUTHORS      Douglas,S.E. and Penny,S.L.
TITLE        The plastid genome of the cryptophyte alga, Guillardia theta:
JOURNAL      complete sequence and conserved synteny groups confirm its common
AUTHORS      ancestry with red algae
TITLE        J. Mol. Evol. 48 (2), 236-244 (1999)
JOURNAL      15 (bases 1 to 121524)
MEDLINE      Douglas,S.E.
REFERENCE    15 (bases 1 to 121524)
AUTHORS      Douglas,S.E.
TITLE        Direct Submission
JOURNAL      Submitted (08-JAN-1998) Institute for Marine Biosciences, National
AUTHORS      Research Council, 1411 Oxford Street, Halifax, Nova Scotia B3H 3Z1,
TITLE        Canada.
COMMENT      On Sep 15, 1998 this sequence version replaced gi:11396 gi:11297
              gi:18103 gi:18281 gi:11383 gi:11407 gi:12539 gi:336730 gi:113300
              gi:338949 gi:2661180.
FEATURES
    source          Location/Qualifiers
                    1..121524
                        /organism="Guillardia theta"
                        /organism="plastid:chloroplast"
                        /db_xref="taxon:55529"
                        /note="plastid"
    tRNA            complement(5..86)
                        /gene="trnY(gua)"
                        /note="codon recognized: CAU"
                        /product="tRNA-Tyr"
    gene            /anticodon=(pos:50..52,aa:Tyr)
                        complement(5..86)
                        /gene="trnY(gua)"
                        complement(181..573)
                        /gene="trp119"
                        complement(181..573)
    CDS             complement(181..573)
                        /gene="trp119"
                        /codon_start=1
                        /product="ribosomal protein L19"
                        /protein_id="AAC35594.1"
                        /db_xref="GI:3602933"
                        /translation="METILILEFVFNKSKNLIQLVESNHNKTLPOLQVGTPLG
                        VSTTEGKRIQATVEGVIYSKSGCINLTITIRVILQGIETERYVILHSEKLSITEL
                        RKSIRAKLIYIRTRSGKATRIQOREN"
    gene            729..3188
                        /gene="clpC"
                        729..3188
                        /gene="clpC"
                        /codon_start=1

```

```

/product="Clp protease amp binding subunit"
/protein_id="AAC35595.1"
/db_xref="GI:3602934"
/translation="MEPEPEKAIVKIMLAQEARLIGHNFGTEQILGLIGEGTGI
AAKVLKSMGVNKKDAVEVEKIIIGSGSGVAEIPPTPAKRVLESLAEANQOLHNY
IGTEHLLGLIRGEGBVAARVLENALDLTKVATVIRLGPDAEVSANQORGTPT
TLEPFGSNLTOKRABEKLDPVDRGROKREIRVIOIIGRRKNNPILIGEGVCKTAAE
GLAQRINNDVDPDILEDKRVYVTLIDGLVAGTKYRGEFEERLKTIDEIRVANNVLY
IDEVHTLIGAGAEGALIDANILKPALARGEQICGATLLEERKHIERDSALERRFO
PVMGSEVPEETIEILYGLDRYERKHKKVIVIDEALSAAKADQYIADRFLEKADID
LIDEGSRVRLNNSOLPPPAARELDELRILKQKDEAVSODPEYAGOLRDEMELKA
GTAAIASHSKGDEENETGSEVTEEDIAQIVAMTGIIPVNNKRSSEKLQWKEEHL
GRIGODEAVVAVSKAIRARVGIKPNPISLIFSGPGTKETELTALASPEFS
ERAMVRIDMSYEMERTVYKSLGSPGYGYNEGGOLTESVRRPTVVLFEIEKGH
PQVFNLLQILDEGRITDSKGRVDEKNTLLITSVGSKVIEKGGGJGFDLSEDT
ESQYGRKALVNEELKQYRPEFLNLDLIVFEROLTKDEVEIAEIMLEKVEPTRISE
KGIQLEVFARFTKLINEGYNPIYGARPLRAVVRMLLEDLISEFLAEKIKEGDZAVY
DVDDGKRVVLIGEKLELVAN"
3316..3388
/gene="trnF(gaa)"
/note="codon recognized: CUU"
/product="tRNA-Phe"
/anticodon=(pos:3349..3351,aa:Phe)
3316..3388
/gene="trnF(gaa)"
complement(3402..5201)
/gene="dnaB"
complement(3402..5201)
/gene="dnaB"
/codon_start=1
/translation="except=(pos:complement(5199..5201),aa:Met)
/product="replication helicase subunit"
/protein_id="AAC35596.1"
/db_xref="GI:3602935"
/translation="MKVTEFVNLKAEKHLIQLILVEKDFLOLYTKIDTQVEYFQSHO
IIVRSNITLHTKKVLINLNLISLHSEVSEVIVLITLFDNRNPNLINDSOL
VLIDNFIIRRELONSCAKIVISLPNPELSEVTLQKRNLLISNINNYKRVSLISL
LLETILIEDKRTNRSTHYLGFEDDHILVGLQKSLIITIAGPSKGTAFKLSIQRN
VADIDSEFVNIVLSLEMSKQILYRILISNTNTATSRLSREGNISISWEILNRMTLS
NNIYIADDPENLIDVIDIQKSLAQIYDIDIGLIVDIYQLQIKKSSQRIHLSYI
TRYKLIADFNPLVYLSQLSRNVFRLNKRPIISDLKESGCLSYATNOPYELKSDN
VNSKSLTSKVSNNHYLSATLLELIPQVNRVYPIVSLIKRELQNGCYKVVYELDEFIS
VIVSYTEHVYVILNGKRIILELTVDLVATLDIQVLIYNNTEVDLSSNVITSSVNL
ICMNRINVTDEMTLPKTNFENALVHNSIEDQADVLPYLRDSTYIQHNLKQSDIMDC
EVIYAKHRGTIGMVLNIPNPNVSEFNLIKES"
3594..4073
/note="signal element intein: encodes putative intein
spliced out of the replication helicase subunit"
3561..5433
/gene="trnW(cca)"
/note="codon recognized: GGU"
/product="tRNA-Trp"
/anticodon=(pos:5394..5396,aa:Trp)
5361..5433
/gene="trnW(cca)"
5588..6013
/gene="trnW(cca)"
5588..6013
/gene="trp111"
/gene="trp111"
/codon_start=1
/product="ribosomal protein L11"
/protein_id="AAC35597.1"
/db_xref="GI:3602936"
/translation="MAKKVVAIIRIKLALLEGKATPAPVGPALGORGVNIYFCKQVNA
KRAEDKAGLITPEIVETIYEDKSYTFVKTPTPASYLAKAAGVQKSGSNPKKTVQGSYTK
KQVEELAQKRLDNLNRLRESAIRITEGAKNMGIVTD"
6033..6737
/gene="trp11"
6033..6737
/gene="trp11"
/gene="trp11"

```

Query Match 72.5%; Score 17.4; DB 8; Length 121524;
Best Local Similarity 94.7%; Pred. No. 2.4e+02;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 tcgtggaagcttcgtaac 19
 Db 11476 TCGTCGACGCTTCTTAAAC 11494

RESULT 8

LOCUS

AF007218

5905 bp DNA

VRT

14-JUL-1997

DEFINITION

Tetradodon fluvialtilis transcription factor snf5 (SNF5PF) gene,
 complete cds.

ACCESSION

AF007218

VERSION

AF007218.1 GI:2253408

KEYWORDS

SOURCE

Tetradodon fluvialtilis.

ORGANISM

Tetradodon fluvialtilis

REFERENCE

Yao, C.W.,

Liu, J.H. and Huang, C.J.

Genomic structure and sequence of SNF5 gene of puffer fish

AUTHORS

Yao, C.W.,

Liu, J.H. and Huang, C.J.

Genomic structure and sequence of SNF5 gene of puffer fish

TITLE

Unpublished

JOURNAL

2 (bases 1 to 5905)

REFERENCE

Yao, C.W.,

Liu, J.H. and Huang, C.J.

Genomic structure and sequence of SNF5 gene of puffer fish

AUTHORS

Yao, C.W.,

Liu, J.H. and Huang, C.J.

Genomic structure and sequence of SNF5 gene of puffer fish

TITLE

Unpublished

JOURNAL

2 (bases 1 to 5905)

REFERENCE

Yao, C.W.,

Liu, J.H. and Huang, C.J.

Genomic structure and sequence of SNF5 gene of puffer fish

AUTHORS

Yao, C.W.,

Liu, J.H. and Huang, C.J.

Genomic structure and sequence of SNF5 gene of puffer fish

TITLE

Unpublished

JOURNAL

2 (bases 1 to 5905)

REFERENCE

Yao, C.W.,

Liu, J.H. and Huang, C.J.

Genomic structure and sequence of SNF5 gene of puffer fish

AUTHORS

Yao, C.W.,

Liu, J.H. and Huang, C.J.

Genomic structure and sequence of SNF5 gene of puffer fish

TITLE

Unpublished

JOURNAL

2 (bases 1 to 5905)

REFERENCE

Yao, C.W.,

Liu, J.H. and Huang, C.J.

Genomic structure and sequence of SNF5 gene of puffer fish

AUTHORS

Yao, C.W.,

Liu, J.H. and Huang, C.J.

Genomic structure and sequence of SNF5 gene of puffer fish

TITLE

Unpublished

JOURNAL

2 (bases 1 to 5905)

REFERENCE

Yao, C.W.,

Liu, J.H. and Huang, C.J.

Genomic structure and sequence of SNF5 gene of puffer fish

AUTHORS

Yao, C.W.,

Liu, J.H. and Huang, C.J.

Genomic structure and sequence of SNF5 gene of puffer fish

TITLE

Unpublished

JOURNAL

2 (bases 1 to 5905)

REFERENCE

Yao, C.W.,

Liu, J.H. and Huang, C.J.

Genomic structure and sequence of SNF5 gene of puffer fish

AUTHORS

Yao, C.W.,

Liu, J.H. and Huang, C.J.

Genomic structure and sequence of SNF5 gene of puffer fish

TITLE

Unpublished

JOURNAL

2 (bases 1 to 5905)

REFERENCE

Yao, C.W.,

Liu, J.H. and Huang, C.J.

Genomic structure and sequence of SNF5 gene of puffer fish

AUTHORS

Yao, C.W.,

Liu, J.H. and Huang, C.J.

Genomic structure and sequence of SNF5 gene of puffer fish

TITLE

Unpublished

JOURNAL

2 (bases 1 to 5905)

REFERENCE

Yao, C.W.,

Liu, J.H. and Huang, C.J.

Genomic structure and sequence of SNF5 gene of puffer fish

AUTHORS

Yao, C.W.,

Liu, J.H. and Huang, C.J.

Genomic structure and sequence of SNF5 gene of puffer fish

TITLE

Unpublished

JOURNAL

2 (bases 1 to 5905)

REFERENCE

Yao, C.W.,

Liu, J.H. and Huang, C.J.

Genomic structure and sequence of SNF5 gene of puffer fish

AUTHORS

Yao, C.W.,

Liu, J.H. and Huang, C.J.

Genomic structure and sequence of SNF5 gene of puffer fish

TITLE

Unpublished

JOURNAL

2 (bases 1 to 5905)

REFERENCE

Yao, C.W.,

Liu, J.H. and Huang, C.J.

Genomic structure and sequence of SNF5 gene of puffer fish

AUTHORS

Yao, C.W.,

Liu, J.H. and Huang, C.J.

Genomic structure and sequence of SNF5 gene of puffer fish

TITLE

Unpublished

JOURNAL

2 (bases 1 to 5905)

REFERENCE

Yao, C.W.,

Liu, J.H. and Huang, C.J.

Genomic structure and sequence of SNF5 gene of puffer fish

AUTHORS

Yao, C.W.,

Liu, J.H. and Huang, C.J.

Genomic structure and sequence of SNF5 gene of puffer fish

TITLE

Unpublished

JOURNAL

2 (bases 1 to 5905)

REFERENCE

Yao, C.W.,

Liu, J.H. and Huang, C.J.

Genomic structure and sequence of SNF5 gene of puffer fish

AUTHORS

Yao, C.W.,

Liu, J.H. and Huang, C.J.

Genomic structure and sequence of SNF5 gene of puffer fish

TITLE

Unpublished

JOURNAL

2 (bases 1 to 5905)

REFERENCE

Yao, C.W.,

Liu, J.H. and Huang, C.J.

Genomic structure and sequence of SNF5 gene of puffer fish

AUTHORS

Yao, C.W.,

Liu, J.H. and Huang, C.J.

Genomic structure and sequence of SNF5 gene of puffer fish

PROGRESS **, in ordered pieces.

ACCESSION

AL513003

VERSION

AL513003.1 GI:12329144

KEYWORDS

HTG; HTGS; PHASE2.

SOURCE

Oryza sativa.

ORGANISM

Oryza sativa

REFERENCE

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

AUTHORS

Salse, J.,

Cholme, N., Orjeda, G., Regad, F., Lortieux, M., Cooke, R.,

TITLE

1 (bases 1 to 183317)

JOURNAL

2 (bases 1 to 183317)

REFERENCE

Salse, J.,

Cholme, N., Orjeda, G., Regad, F., Lortieux, M., Cooke, R.,

AUTHORS

Salse, J.,

Cholme, N., Orjeda, G., Regad, F., Lortieux, M., Cooke, R.,

TITLE

1 (bases 1 to 183317)

JOURNAL

2 (bases 1 to 183317)

REFERENCE

Salse, J.,

Cholme, N., Orjeda, G., Regad, F., Lortieux, M., Cooke, R.,

AUTHORS

Salse, J.,

Cholme, N., Orjeda, G., Regad, F., Lortieux, M., Cooke, R.,

TITLE

1 (bases 1 to 183317)

JOURNAL

2 (bases 1 to 183317)

REFERENCE

Salse, J.,

Cholme, N., Orjeda, G., Regad, F., Lortieux, M., Cooke, R.,

AUTHORS

Salse, J.,

Cholme, N., Orjeda, G., Regad, F., Lortieux, M., Cooke, R.,

TITLE

1 (bases 1 to 183317)

JOURNAL

2 (bases 1 to 183317)

REFERENCE

Salse, J.,

Cholme, N., Orjeda, G., Regad, F., Lortieux, M., Cooke, R.,

AUTHORS

Salse, J.,

Cholme, N., Orjeda, G., Regad, F., Lortieux, M., Cooke, R.,

TITLE

1 (bases 1 to 183317)

JOURNAL

2 (bases 1 to 183317)

REFERENCE

Salse, J.,

Cholme, N., Orjeda, G., Regad, F., Lortieux, M., Cooke, R.,

AUTHORS

Salse, J.,

Cholme, N., Orjeda, G., Regad, F., Lortieux, M., Cooke, R.,

TITLE

1 (bases 1 to 183317)

JOURNAL

2 (bases 1 to 183317)

REFERENCE

Salse, J.,

Cholme, N., Orjeda, G., Regad, F., Lortieux, M., Cooke, R.,

AUTHORS

Salse, J.,

Cholme, N., Orjeda, G., Regad, F., Lortieux, M., Cooke, R.,

TITLE

1 (bases 1 to 183317)

JOURNAL

2 (bases 1 to 183317)

REFERENCE

Salse, J.,

Cholme, N., Orjeda, G., Regad, F., Lortieux, M., Cooke, R.,

AUTHORS

Salse, J.,

Cholme, N., Orjeda, G., Regad, F., Lortieux, M., Cooke, R.,

TITLE

1 (bases 1 to 183317)

JOURNAL

2 (bases 1 to 183317)

REFERENCE

Salse, J.,

Cholme, N., Orjeda, G., Regad, F., Lortieux, M., Cooke, R.,

AUTHORS

Salse, J.,

Cholme, N., Orjeda, G., Regad, F., Lortieux, M., Cooke, R.,

TITLE

1 (bases 1 to 183317)

JOURNAL

2 (bases 1 to 183317)

REFERENCE

Salse, J.,

Cholme, N., Orjeda, G., Regad, F., Lortieux, M., Cooke, R.,

AUTHORS

Salse, J.,

Cholme, N., Orjeda, G., Regad, F., Lortieux, M., Cooke, R.,

TITLE

1 (bases 1 to 183317)

JOURNAL

2 (bases 1 to 183317)

REFERENCE

Salse, J.,

Cholme, N., Orjeda, G., Regad, F., Lortieux, M., Cooke, R.,

AUTHORS

Salse, J.,

Cholme, N., Orjeda, G., Regad, F., Lortieux, M., Cooke, R.,

TITLE

1 (bases 1 to 183317)

JOURNAL

2 (bases 1 to 183317)

REFERENCE

Salse, J.,

Cholme, N., Orjeda, G., Regad, F., Lortieux, M., Cooke, R.,

AUTHORS

Salse, J.,

Cholme, N., Orjeda, G., Regad, F., Lortieux, M., Cooke, R.,

TITLE

1 (bases 1 to 183317)

JOURNAL

2 (bases 1 to 183317)

REFERENCE

Salse, J.,

Cholme, N., Orjeda, G., Regad, F., Lortieux, M., Cooke, R.,

AUTHORS

Salse, J.,

TITLE Direct Submission
JOURNAL Submitted (10-JAN-2001) Genomics, Monsanto, 800 North Lindbergh Blvd., Creve Coeur, MO 63167, USA
COMMENT Derived from rice genomic sequences generated from the Monsanto Rice Genome Sequencing project. Please see <http://www.rice-research.org> for more information. The sequence data were produced primarily in the laboratories of Dr. Leroy Hood at the University of Washington in Seattle.

FEATURES
SOURCE 1. .224
 /organism="Oryza sativa"
 /db_xref="taxon:4530"
 repeat_region 1. .224
 /note="microsatellite MRG6352"
 /rpt_type="tandem"
 /rpt_unit="cgtt"
BASE COUNT 30 a 57 c 42 g 95 t
ORIGIN

Query Match 70.8%; Score 17; DB 8; Length 224;
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 acgttcgttaacgttcg 24
 ||||||||||||||||
 Db 50 ACgttcGTTAAGtTCG 66

RESULT 11
LOCUS AP003370/c
DEFINITION Oryza sativa chromosome 1 clone B1129H01, *** SEQUENCING IN PROGRESS ***
ACCESSION AP003370
VERSION AP003370.1 GI:13365587
KEYWORDS HTG; HTGS; PHASE2
SOURCE Oryza sativa (cultivar: Nipponbare) DNA, clone: B1129H01.
ORGANISM Oryza sativa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE 1 (sites)
AUTHORS Sasaki, T., Matsumoto, T. and Yamamoto, K.
TITLE Oryza sativa nipponbare (Ga3) genomic DNA, chromosome 1, BAC clone: B1129H01
JOURNAL Published Only in Database (2001) In press
REFERENCE 2 (bases 1 to 101803)
AUTHORS Sasaki, T., Matsumoto, T. and Yamamoto, K.
JOURNAL Direct Submission
 Submitted (07-MAR-2001) Takuji Sasaki, National Institute of Agricultural Resources, Rice Genome Research Program; Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
 (E-mail: tsasaki@affrc.go.jp, URL: <http://rgp.dna.affrc.go.jp/>, Tel: 81-298-38-7441, Fax: 81-298-38-7468)
COMMENT NOTE: It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have provided by the submitter. This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.
 * NOTE: This is a 'working draft' sequence.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

FEATURES
SOURCE 1. 101803
 /organism="Oryza sativa"
 /cultivar="Nipponbare"
 /db_xref="taxon:4530"
 /chromosome="1"
 /clone="B1129H01"
BASE COUNT 29451 a 21109 c 21813 g 29380 t 50 others

ORIGIN

Query Match 70.8%; Score 17; DB 2; Length 101803;
 Best Local Similarity 100.0%; Pred. No. 4e+02;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 acgttcgttaacgttcg 24
 ||||||||||||||||
 Db 96635 ACgttcGTTAAGtTCG 96619

RESULT 12
LOCUS AP003987
DEFINITION Oryza sativa chromosome 2 clone OJ1057_A12, *** SEQUENCING IN PROGRESS ***
ACCESSION AP003987
VERSION AP003987.1 GI:15076790
KEYWORDS HTG; HTGS; PHASE2
SOURCE Oryza sativa (cultivar: Nipponbare) DNA, clone: OJ1057_A12.
ORGANISM Oryza sativa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 104718)
AUTHORS Sasaki, T., Matsumoto, T. and Yamamoto, K.
TITLE Oryza sativa nipponbare (Ga3) genomic DNA, chromosome 2, BAC clone: OJ1057_A12
JOURNAL Published Only in Database (2001) In press
REFERENCE 2 (bases 1 to 104718)
AUTHORS Sasaki, T., Matsumoto, T. and Yamamoto, K.
JOURNAL Direct Submission
 Submitted (01-AUG-2001) Takuji Sasaki, National Institute of Agricultural Resources, Rice Genome Research Program; Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
 (E-mail: tsasaki@affrc.go.jp, URL: <http://rgp.dna.affrc.go.jp/>, Tel: 81-298-38-7441, Fax: 81-298-38-7468)
COMMENT NOTE: It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have provided by the submitter. This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.
 * NOTE: This is a 'working draft' sequence.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

FEATURES
SOURCE 1. 104718
 /organism="Oryza sativa"
 /cultivar="Nipponbare"
 /db_xref="taxon:4530"
 /chromosome="2"
 /clone="OJ1057_A12"
BASE COUNT 31132 a 21354 c 21209 g 30922 t 101 others

ORIGIN

Query Match 70.8%; Score 17; DB 2; Length 104718;
 Best Local Similarity 100.0%; Pred. No. 4e+02;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 acgttcgttaacgttcg 24
 ||||||||||||||||
 Db 79837 ACgttcGTTAAGtTCG 79853

RESULT 13
LOCUS AP003196
DEFINITION AP003196 149498 bp DNA HTG 21-FEB-2001

DEFINITION Oryza sativa chromosome 1 clone B1008C01. *** SEQUENCING IN PROGRESS ***. In ordered pieces.

ACCESSION AP003196

VERSION AP003196.1 GI:13027226

KEYWORDS HTG; HTGS; PHASE2.

SOURCE Oryza sativa (cultivar:Nipponbare) DNA, clone:B1008C01.

ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE 1 (sites)

AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.

TITLE Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, BAC clone:B1008C01

JOURNAL Published Only In Database (2001) In press

REFERENCE 2 (bases 1 to 149498)

AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.

TITLE Direct Submission

JOURNAL Submitted (19-FEB-2001) Takuji Sasaki, National Institute of Agricultural Resources, Rice Genome Research Program; Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan (E-mail:tsasaki@ab.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/, Tel:81-298-38-7441, Fax:81-298-38-7468)

COMMENT NOTE: It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have provided by the submitter. This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.

* NOTE: This is a 'working draft' sequence.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

FEATURES

source

1..149498

/organism="Oryza sativa"

/cultivar="Nipponbare"

/db_xref="taxon:4530"

/chromosome="1"

/clone="B1008C01"

BASE COUNT 41722 a 33641 c 31241 g 42644 t 250 others

ORIGIN

Query Match 70.8%; Score 17; DB 2; Length 149498;

Best Local Similarity 100.0%; Pred. No. 4e+02;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 8 acgttcgtaacgttcg 24

|||||

Db 25666 ACGTTCGTTAACGTTTCG 25682

RESULT 14

AP003074/c 150379 bp DNA PLN 29-MAR-2001

LOCUS Oryza sativa genomic DNA, chromosome 1, BAC clone:OSJNBa0004G10.

DEFINITION AP003074

ACCESSION AP003074

VERSION AP003074.2 GI:13466822

KEYWORDS

SOURCE Oryza sativa (cultivar:Nipponbare) DNA, clone:OSJNBa0004G10.

ORGANISM Oryza sativa

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE 1 (sites)

AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.

TITLE Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, BAC clone:OSJNBa0004G10

JOURNAL Published Only In Database (2001) In press

REFERENCE 2 (bases 1 to 150379)

AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.

TITLE Direct Submission

JOURNAL Submitted (10-JAN-2001) Takuji Sasaki, National Institute of Agricultural Resources, Rice Genome Research Program; Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan (E-mail:tsasaki@ab.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/, Tel:81-298-38-7441, Fax:81-298-38-7468)

COMMENT On Mar 28, 2001 this sequence version replaced gi:12082349. Genes were predicted from the integrated results of the following: GENSCAN1.0, BLASTX2.0, BLASTX2.0 as well as SplicePredictor (October 1998 version). The genomic sequence was searched against NCBI NonRedundant Protein database, or (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at RGP. Protein homologues of the coding regions were searched against NCBI NonRedundant Protein database with BLASTP2.0. ESTs represent the identified cDNA sequences using BLASTN 2.0 with the corresponding DBE accession no. and RGP clone ID.

A gene with identity or significant homology to a protein is classified based on the protein name to indicate the homology level such as same name, 'putative-' and '-like protein'. A gene without significant homology to any protein but with EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted with a gene prediction program is classified as a 'hypothetical' protein.

The orientation of the sequence is from -21M13 to M13rev of the BAC clone. This sequence of OSJNBa0004G10 clone has an overlap with P0034C11 (DBJ: AP002865) clone at the position 78,934 to 150,379 of 3' end. The sequence of this clone ends at the position 71,446 of P0034C11. Detailed information on overlap and assembly quality together with annotation of this entry is available at <http://rgp.dna.affrc.go.jp/Genomeseq.html>.

FEATURES

source

1..150379

/organism="Oryza sativa"

/cultivar="Nipponbare"

/db_xref="taxon:4530"

/chromosome="1"

/clone="OSJNBa0004G10"

join(3892..3894,4077..4235,4302..4379,4831..4989)

/gene="OSJNBa0004G10.1"

join(3892..3894,4077..4235,4302..4379,4831..4989)

/gene="OSJNBa0004G10.1"

/note="hypothetical protein"

/codon_start=1

/protein_id="BAB40054.1"

/db_xref="GI:13466823"

/translation="MTYRPMCFTEWDMENPNONGTRAYRNEFRSDYLRRKDEHRRIAEALAEHYIRSVRRRTGRCFVCPENIYDDDFVGNPRGSGTTRMIDNTPSYHGO KITESEYQVQRKLDHNAHSDRPRGR"

complement(join(6294..6572,7027..7047))

/gene="OSJNBa0004G10.2"

complement(join(6294..6572,7027..7047))

/gene="OSJNBa0004G10.2"

complement(join(6294..6572,7027..7047))

/gene="OSJNBa0004G10.2"

/note="hypothetical protein"

/codon_start=1

/protein_id="BAB40055.1"

/db_xref="GI:13466824"

/translation="MIPSRPQSLVDDGGGGGGRARORVGDGRGGGSVDGGRTPRWRWRRTATTAVDDGGGGGGRVRRORGMIALRGAIPGDSIPSSILV"

join(9066..9225,10554..10996)

/gene="OSJNBa0004G10.3"

join(9066..9225,10554..10996)

/gene="OSJNBa0004G10.3"

/note="hypothetical protein"

/codon_start=1

/protein_id="BAB40056.1"

/db_xref="GI:13466825"

/translation="MEENEVEDNIPDPAQVGFEGNGTEEEDADGNDVADLGOI LQVKECEPLSPISLISVAAGVADADVAEDSAPRRRGTTTRDATT PPRHGAAPRLRRTPPPNAAAPRAAPRCPTRCARHRAAAAPRRHSAAT PALRCRVRIRIVNEMVNEVNEVNERELNVERA"

complement(join(13279..13346,13549..14688,14843..14926,15165..15308,15344..15407))

/gene="OSJNBa0004G10.4"

complement(join(13279..13346,13549..14688,14843..14926,

CDS

gene

15165. 15308.15344. 15407))
 /gene="OSJNBa0004G10.4"
 /note="hypothetical protein
 similar to Oryza sativa OSJNBa0049B20.1"
 /codon_start=1
 /protein_id="BAB40057.1"
 /db_xref="GI:13486826"
 /translation="MPDWIGACVADGQTPVRCARCAPSSRVVFPAGNRTSEFNGA
 LASIDSPARLSMSRLAPSTRYSPRNAPARAIAGNSTPSRRYKPRINGCTEPMSA
 PMKEVDPVADKRMQKDKGRWVLLACVPHVIRGDEAGADNVYKAALAPRATRY
 VHRSTAPRRKTIIDHFPVAGADCHRRLLLYASQGEPEEPVLDGTRPLGHEHFPK
 AFICDTLKHSTRLPDHGFPLHFGNGLVAITKTIHVADLHIVGSKRAALITF
 SVPEIMATVVDYPPRDRMGNGVVAHKTIVMWVDSYGLSCDISARONLREVP
 PPDELPGTDLDRKRCVGRGDLRYMGJHEREYDGRPVVSMWTLVODDAGTWRLD
 COPFLKIDNDEGKATKLPREIPVAFIHPGLPGNVAFEFMRSLFGVINYTRKYLE
 MOFLMLNPMHSSRFVRAACPNGLIMHALCEISTRNSEFMNG"
 complement(19438. 20550)
 /gene="OSJNBa0004G10.5"
 complement(19438. 20550)
 /gene="OSJNBa0004G10.5"
 /note="hypothetical protein
 similar to Oryza sativa OSJNBa0049B20.1"
 /codon_start=1
 /protein_id="BAB40058.1"
 /db_xref="GI:13486827"
 /translation="MQREAAAGMWVILVSPVHAHDEFHPPGTETLRINFGAPFAS
 RTVPRRIAPDRKAIIDNYPLAADERHGRLLVATOGDPDEPPALDAFYLRPLGVH
 HGFAYAYICDTTDEASRLPDHGFALIVGNVGLCSISFVVALEOPAPASGA
 TLLIYRSDSPADDELSYPPHDPWNGVNGVSHODRLIMWVLSYGLITCDVYVDDP
 PDHVPYLPDSELPAGTPDLERKRCVGSAGRLKLYVIDDEPDDPIYRMTLLDED
 AGEWGFDCASVVAIWDEAVKATKLPQVPAVALIHPTGPDVVYEFRLSRITFADV
 RARLLERFEEMLHPMPRHSSQFVRAMNLETLQ"
 complement(join(22493. 22509,23300. 23459,23521. 23766,
 24024. 24080))
 /gene="OSJNBa0004G10.6"
 complement(join(22493. 22509,23300. 23459,23521. 23766,
 24024. 24080))
 /gene="OSJNBa0004G10.6"
 /note="hypothetical protein"
 /codon_start=1
 /protein_id="BAB40059.1"
 /db_xref="GI:13486828"
 /translation="MSCRAGTSGSAPPYQPPPLIPRAGRGGRHROAPRCIKAS
 AGNKRIVQIHGPDEQVYITWTLHDRVHAKRREIYEPFVETWQGRKTRSEMRKG
 QLMIGABGSGVRLRDPTKTPSVGLPPEFVMEVPSLPRTVPEVQHDNLN"
 join(26535. 27327,27406. 27422)
 /gene="OSJNBa0004G10.7"
 join(26535. 27327,27406. 27422)
 /gene="OSJNBa0004G10.7"
 /note="hypothetical protein"
 /codon_start=1
 /protein_id="BAB40060.1"
 /db_xref="GI:13486829"
 /translation="MANPVPCPSAAVAGPGGCDGVGSSSNVNAHVRHNGVYRL
 CSSCVILGRDAFCYCLDVPFGDAPFODDFNPIYSCCCEVEPYAAHNLCTIDPS
 IFVCPACAAAGRTTYAPSSGAPALAEKVLVAAARLHSAVAAAALAEQEDKL
 IREAAARARADVDVACRVLEAEARDAKEAALAPSVLTKTPKNSANRSRDK
 PLKINSICPAPLAFAAAAAASSTPLSTPSPAGEKPKQGVLLAV"
 complement(join(28267. 28383,28533. 28659,28780. 28926,
 29466. 30857,31383. 31453))
 /gene="OSJNBa0004G10.8"
 complement(join(28267. 28383,28533. 28659,28780. 28926,
 29466. 30857,31383. 31453))
 /gene="OSJNBa0004G10.8"
 /note="hypothetical protein
 similar to Oryza sativa OSJNBa0049B20.1"
 /codon_start=1
 /protein_id="BAB40061.1"
 /db_xref="GI:13486830"
 /translation="MIRSRGSGTGTSPVPCARAANRDRGKPIEGSSPPFVHTPIPIA
 TSHSNAAAAAASEVRSKRMEHSHPLDGAAPFRDERMAAPPRATVILACVP
 SVSSDSDEAGDHIAFDWRDPGVSLTLRSDSVSPAPADPCPDHDDHRYVAV
 DSAAGILLGARSADHPGCVALGPDPAISCTNCGYILICHTALITLMAVILPPCSDEY

RLCAGNVGMIRTRADRDHPIRELLAEQIESGNCIHRATILIRYHEJGMASTYVNP
 PGRSMCGDVIYHAGMLMWVDSFGTLTCVFAKPMRVRVPEGGKLPYSSDADH
 AKHCNVANSDEGLAFVQIHVDYTAAGRAPSTIMSMITLOOSAGAESVSLHHRV
 VDEIMDHYTYRKTAMPREPVIALHHPREGLVPEFQETSNRSMWAFVADLTRVLEC
 KKYKMPOLPMYHSSRRAWEIHPISICRGDDEPDTIACVIONLDELIDNPSD
 KADELSTGTGLFINPRQELRNATAFKYSFIYVATDYEALCLMRNVSHVSMO
 GSNIVPVSASISLISGLTCLTQVFNWMMADDIVRDLGTLSSNRITDG"
 complement(join(32717. 32742,33198. 3313,33434. 33558,
 33802. 33841,34125. 35252,36069. 36148,36260. 36304))
 /gene="OSJNBa0004G10.9"
 complement(join(32717. 32742,33198. 3313,33434. 33558,
 33802. 33841,34125. 35252,36069. 36148,36260. 36304))
 /gene="OSJNBa0004G10.9"
 /note="hypothetical protein
 similar to Oryza sativa OSJNBa0049B20.1"
 /codon_start=1
 /protein_id="BAB40062.1"
 /db_xref="GI:13486831"
 /translation="WSSSSAATRCGPRRTAKIDTSVFGESSINSFAIYGADK
 LIDGSPPEVHTPTPIATDHSNMAAAASEVRSKRMEHSPHLPDGAAPFERDE
 RMAAPREATVILACVPSVSSDDEPAGHIAFDMDPDGVSLTLRQSDSVFSP
 APDPCPRDDHPYVAVDSAGGLLRGARSADHFGVAGLGFDPAPCANDGCIYL
 CHATIRMAVIXPCSDERYRLIYAGVGMIRTRADRDHPIRELLAEQIESGNCIHRAT
 LLRSQELGLGMASTKYNIPGRSKCGDGYIYHAGMLMWVDSFGTLTCAPFAK
 DMREVPLEGGKLPYSSDADHAKHRCVNSOGEIAFVQIHVDYTAAGRAPSTIM
 WTLQSDAGESSVSLRHRVRDEIMDHYTYRKTAMPREPVIALHHPREGLVPEFQ
 ITSRSMWAFVADLTRVLECXYKMPOLPMYHSSRRAWEIHPISICRGDDETDG
 TSYNVMAFADVRLNDELINFRSRADELSTGTGLFINPRQELRNATAFKYL
 SFVIKATDAEASCLMDVFSVSMDSNIVPRSEYEDDD"
 join(36644. 36865,37413. 37473,37564. 37949)
 /gene="OSJNBa0004G10.10"
 join(36644. 36865,37413. 37473,37564. 37949)
 /gene="OSJNBa0004G10.10"
 /note="hypothetical protein
 similar to Oryza sativa OSJNBa0049B20.2"
 /codon_start=1

Query Match 70.8%; Score 17; DB 8; Length 150379;
 Best Local Similarity 100.0%; Pred. No. 4e+2;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 8 acgttgtaacgttcg 24
 DB 11078 ACgTtGtTAACgTTCG 11062

RESULT 15
 AP003211/C
 LOCUS
 DEFINITION
 Oryza sativa chromosome 1 clone OSJNBa0011P19, *** SEQUENCING IN
 PROGRESS ***, in ordered pieces.
 ACCESSION
 AP003211
 VERSION
 AP003211.1 GI:13027241
 KEYWORDS
 HTG; HTGS; PHASE2
 SOURCE
 Oryza sativa (cultivar: Nipponbare) DNA, clone: OSJNBa0011P19.
 ORGANISM
 Oryza sativa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.
 1 (sites)
 REFERENCES
 Sasaki,T., Matsumoto,T. and Yamamoto,K.
 Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, BAC
 clone:OSJNBa0011P19
 Published Only in DataBase (2001) In press
 2 (bases 1 to 150641)
 Sasaki,T., Matsumoto,T. and Yamamoto,K.
 Direct Submision
 Submitted (19-FEB-2001) Takuji Sasaki, National Institute of
 Agricultural Resources, Rice Genome Research Program, Kannondai
 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
 (E-mail:tsasaki@affrc.go.jp, URL:http://rpg.dna.affrc.go.jp/,
 Tel:81-298-38-7441, Fax:81-298-38-7468)
 NOTE: It currently consists of 1 contigs. Gaps between the contigs

COMMENT

are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have provided by the submitter. This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.

* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

FEATURES:

source

1. 150641

/organism="Oryza sativa"

/cultivar="Nipponbare"

/db_xref="taxon:4530"

/chromosome="1"

/clone="OSJNBa0011P19"

BASE COUNT 41696 a 33905 c 33240 g 41500 t 300 others

ORIGIN

Query Match

70.8%; Score 17; DB 2; Length 150641;

Best Local Similarity 100.0%; Pred. No. 4e+02;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 acgttcgttaacgttcg 24

Db 148877 ACGTCGTTACGTCG 148861

Search completed: March 11, 2002, 20:56:28
Job time: 4346 sec